

MAT

Monografia/Est
G939e

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
DEPARTAMENTO DE ESTATÍSTICA

ENSAIOS DE BIOEQUIVALÊNCIA: DO PLANEJAMENTO À ANÁLISE

Autor: Katrine Guewehr

Orientador: Prof. Dr. João Riboldi

Trabalho de conclusão apresentado para a obtenção
do grau de Bacharel em Estatística

Porto Alegre, julho de 2004.

AGRADECIMENTOS

Agradeço, em primeiro lugar a Deus, por me ajudar a chegar até aqui...

Agradeço a todos àqueles que de alguma maneira contribuíram para a realização deste trabalho, me apoiando sempre que necessário.

Agradeço em especial ao professor, orientador e acima de tudo amigo, Prof. João Riboldi, pelo carinho, atenção e apoio em todas as horas, desde as cadeiras de planejamento de experimentos.

Ao professor Álvaro Vigo, a quem dedico grande admiração e carinho, agradeço por tudo.

Aos demais professores pelos conhecimentos transmitidos que serão de grande importância para o meu futuro profissional.

Agradeço também aos professores Bruce B. Duncan, Maria Inês Schmidt e Sotero Serrate Mengue, do Programa de Pós-Graduação em Epidemiologia, por terem me dado a primeira oportunidade de aprendizado prático, através da Bolsa de Iniciação Científica.

Às amigas de todas as horas, Margarete, Patrícia e Gabriela, pelo apoio incondicional.

Aos amigos do curso de Estatística, principalmente aqueles que ingressaram comigo em 2000, em especial Filipe e Raquel.

À minha família, em especial meus pais, Breno e Denise, e meus irmãos Tharlany e Breno Neto, por me darem o carinho e o apoio necessários para enfrentar as dificuldades do dia a dia.

ÍNDICE

1. INTRODUÇÃO	9
2. PLANEJAMENTO	12
2.1 Ensaio Clínicos	12
2.2 Método Bioanalítico	13
2.2.1 Introdução	13
2.2.2 Pré-validação	14
2.2.2.1 Exatidão, Precisão e Recuperação	14
2.2.2.2 Linearidade e Limites de Quantificação	15
2.2.2.3 Especificidade e Seletividade	15
2.2.3 Estabilidade dos Fármacos em Fluidos Biológicos	16
2.2.4 Validação do Método	17
2.3 Planejamento Experimental	17
2.4 Caracterização do Planejamento Experimental	18
2.4.1 Delineamento Experimental	18
2.4.2 Aleatorização	19
2.4.3 Período de Eliminação (“Washout”) e Efeitos Residuais (“Carry-Over Effects”)	19
2.4.4 Número de Voluntários	19
2.4.5 Seleção do Delineamento Experimental	19
2.5 O Planejamento “Crossover”	20
2.6 Tipos de Delineamentos “Crossover”	20
2.6.1 Delineamento “Crossover” para Dois Medicamentos	20
2.6.2 Delineamento “Crossover” para Três Medicamentos	23
2.6.3 Delineamento “Crossover” para Quatro Medicamentos	23
3. ANÁLISE DE ENSAIOS DE BIOEQUIVALÊNCIA	25
3.1 Os Dados	25
3.1.1 Principais Medidas Farmacocinéticas	25
3.1.2 Estudos com Doses Múltiplas	28
3.1.3 Outras Considerações Importantes	29

3.2 Métodos de Análise de Bioequivalência	30
3.2.1 Introdução	30
3.2.2 Modelo Estatístico	31
3.2.3 Inferência Estatística sobre Efeitos de um Delineamento Crossover 2x2	32
3.2.3.1 Efeitos Residuais	33
3.2.3.2 O Efeito da Droga	35
3.2.3.3 O Efeito de Período	37
3.2.4 Análise de Variância	38
3.3 Métodos Estatísticos para Bioequivalência Média	42
3.3.1 Introdução	42
3.3.2 Intervalos para Bioequivalência	43
3.3.2.1 Medidas Farmacocinéticas sem Transformação	43
3.3.2.2 Medidas Farmacocinéticas com Transformação Logarítmica	44
3.3.3 Avaliação da Bioequivalência Média	45
3.3.3.1 Intervalo de Confiança	46
3.3.3.2 Teste de Hipóteses	49
3.3.4 Poder dos Testes Unilaterais	53
3.3.5 Dimensionamento de Amostra	54
3.4 Procedimentos Bootstrap	56
3.4.1 Introdução	56
3.4.2 Técnica Bootstrap	57
3.4.3 Intervalo de Confiança Bootstrap em Estudos de Bioequivalência	58
3.5 Análise de Medidas Repetidas em Planejamentos Crossover 2x2	60
3.5.1 Caracterização	60
3.5.2 Estrutura de Covariância	61
3.5.3 Metodologia de Modelos Mistos para Análise de Medidas Repetidas de Dados em Estudos Crossover	62
3.6 Algumas Considerações sobre Bioequivalência	64
3.6.1 Observações Atípicas (“Outliers”) em Estudos de Bioequivalência	64
3.6.2 Uso de Covariáveis	67
3.6.3 Dados Faltantes (“Missings Values”)	67

3.6.4 Bioequivalência Individual e Populacional	71
4. EXEMPLO	73
4.1 Análise Preliminar do Exemplo usando a variável ASC_{t_k}	78
4.1.1 Análise dos Efeitos Residuais	78
4.1.2 Análise do Efeito de Tratamento	80
4.1.3 Análise de Efeito de Período	81
4.1.4 Análise de Variância	82
4.1.5 Intervalo de Confiança para Bioequivalência Média	83
4.1.5.1 Intervalos de Bioequivalência	83
4.1.5.2 Intervalos de Confiança para a Bioequivalência: Método Paramétrico	84
4.1.5.3 Intervalos de Confiança para a Bioequivalência: Método Não-Paramétrico	85
4.2 Análise Preliminar do Exemplo usando a variável C_{max}	86
4.2.1 Análise Descritiva	86
4.2.1.1 Análise Descritiva das Concentrações Plasmáticas	86
4.2.1.2 Análise Descritiva das Medidas Farmacocinéticas	89
4.3 Análise de Variância das Medidas Farmacocinéticas	90
4.3.1 Análise de Variância das Medidas Farmacocinéticas para o Modelo Completo	90
4.3.1.1 Análise de Variância para a variável T_{max}	91
4.3.1.2 Análise de Variância para a variável C_{max}	93
4.3.1.3 Análise de Variância para a variável K_e	96
4.3.1.4 Análise de Variância para a variável $t_{1/2}$	98
4.3.1.5 Análise de Variância para a variável ASC_t	100
4.3.1.6 Análise de Variância para a variável ASC_{inf}	102
4.3.2 Análise de Variância das Medidas Farmacocinéticas para o Modelo Reduzido	105
4.4 Análise de Medidas Repetidas da Variável Concentração Plasmática ao longo Do Tempo	105
5. PROCEDIMENTOS COMPUTACIONAIS DE ANÁLISE	114

6. CONSIDERAÇÕES FINAIS	119
7. REFERÊNCIAS BIBLIOGRÁFICAS	121
8. ANEXOS	124

RESUMO

Os Ensaio de Bioequivalência são caracterizados por planejamentos “crossover”, onde seguindo o protocolo da ANVISA (Agência Nacional da Vigilância Sanitária), no mínimo 24 sujeitos são alocados aleatoriamente para receber as formulações de teste e referência. As medidas farmacocinéticas envolvidas no estudo são as áreas sob a curva de concentração sanguínea do tempo zero ao tempo t e do tempo zero ao infinito, o pico de concentração máxima observada e o tempo que o fármaco e/ou metabólito leva para atingir essa concentração máxima. Adicionalmente devem ser determinados a depuração, o volume aparente da distribuição, a constante de velocidade de absorção do fármaco no organismo e a meia-vida de eliminação do fármaco. Neste trabalho enfoca-se as etapas de planejamento e análise dos ensaios de bioequivalência, contemplando a fase de coleta dos dados, técnicas especializadas de análise (análise de medidas repetidas) e os procedimentos computacionais pertinentes. Este trabalho é extremamente enriquecido pela apresentação de um exemplo, onde são aplicadas todas técnicas descritas a partir de dados hipotéticos.

Palavras Chave: bioequivalência, planejamento crossover, intervalos de bioequivalência, análise de medidas repetidas em planejamento “crossover”.

ABSTRACT

Bioequivalence's Studies are characterized by crossover design 2x2, following ANVISA (Agência Nacional de Vigilância Sanitária); at least 24 subjects are randomly placed to receive the reference's drugs and test's drugs. The pharmacokinetics measures in study are the areas under the curve of plasmatic concentration by time zero to time t and time zero to infinite, peak of observed maximum concentration and the time that the drug leads to reach the maximum concentration. Additionally must be determined the purification, the apparent volume of the distribution, the absorption's speed constant of the drug on organism and the stocking-life of drug's elimination. In this work, we are interested in the design and analysis of bioequivalence's studies, contemplating the data's collection, specialized techniques of analysis (repeated measures) and the pertinent computational procedures. We present this using a extremely rich example, which describes techniques applied by hypothetical data.

Key words: bioequivalence, crossover design, bioequivalence's intervals, repeated measures in crossover design.

1 INTRODUÇÃO

Uma área de interesse crescente na Estatística, ou mais especificamente, na Bioestatística, é a que trata de métodos estatísticos para análise de ensaios de bioequivalência de produtos farmacêuticos. A partir dessa motivação e curiosidade intelectual foi escolhido o assunto para o presente trabalho. O objetivo deste trabalho é abordar o importante e atual tema de Ensaios de Bioequivalência, passando pela coleta de dados e pelas diferentes formas de análise, fazendo as ilustrações numéricas e computacionais pertinentes.

Segundo a Agência Nacional de Vigilância Sanitária (ANVISA), bioequivalência consiste na demonstração de equivalência farmacêutica entre produtos apresentados sob a mesma forma farmacêutica, contendo idêntica composição qualitativa e quantitativa de princípio(s) ativo(s), e que tenham comparável biodisponibilidade (quantidade e velocidade de absorção do fármaco no sistema circulatório), quando estudados sob um mesmo delineamento experimental.

Os estudos de biodisponibilidade iniciaram-se a partir de 1945, apresentando significativo desenvolvimento na década de 1960. Neste período a adoção de técnicas analíticas eficientes possibilitou o desenvolvimento de métodos sensíveis o suficiente para permitir a quantificação de drogas ou metabólitos, inicialmente na urina, e posteriormente no plasma, o que possibilitou a avaliação e comparação da biodisponibilidade de diferentes formulações em voluntários, bem como a demonstração de que diferenças significativas entre estas possam ocorrer. Após a introdução da legislação de registro compulsório de medicamentos em 1969, que facilitou a entrada de medicamentos genéricos no mercado canadense, o diretório canadense do departamento federal de saúde e bem-estar começou a utilizar bioequivalência como uma medida para aprovar o registro de um medicamento durante a década de 1970. Este programa, juntamente com outras informações, foi analisado pelo FDA, órgão a editar as primeiras diretrizes para a realização de estudos de bioequivalência em 1977, nos Estados Unidos. A aplicação destas diretrizes foi ampliada no ato de restauração do termo de patente e competição do preço da droga (“Drug Price Competition and Patent Term Restoration Act”) de 1984, o qual concedia ao FDA poderes para autorizar a aprovação de drogas genéricas sem evidências clínicas de segurança ou eficácia, desde que a droga seja bioequivalente ao produto inovador.

Os estudos de biodisponibilidade e bioequivalência passaram a ser exercidos de forma rotineira no Brasil a partir dos genéricos (Lei nº 9787/99). Neste contexto, destaca-se, pelo seu

pioneirismo a participação do Dr. Gilberto de Nucci, responsável pela implantação da Unidade Miguel Servet no Departamento de Farmacologia da Faculdade de Ciências Médicas da Unicamp.

De acordo com a Resolução RE nº 478, de 19 de março de 2002, publicada no Diário Oficial de 20/03/2002, em Estudos de Bioequivalência para Medicamentos Genéricos, as medidas farmacocinéticas deverão ser obtidas das curvas de concentração sanguínea do fármaco em função do tempo, e analisadas estatisticamente para determinação de bioequivalência.

As medidas farmacocinéticas usadas para medir biodisponibilidade que devem ser determinadas são: i) a área sob a curva de concentração sanguínea em função do tempo, do tempo zero ao tempo t (ASC- t), em que t é o tempo relativo à última concentração determinada experimentalmente; ii) a área sob a curva de concentração sanguínea em função do tempo, calculada do tempo zero ao tempo infinito (ASC-inf); iii) o pico de concentração máxima (C_{max}) do fármaco e/ou metabólito e o tempo para atingir esse pico (T_{max}) que devem ser obtidos diretamente sem interpolação dos dados. Adicionalmente devem ser determinados a depuração (D), o volume aparente de distribuição (V_d) e a meia-vida de eliminação ($t_{1/2}$) do fármaco e/ou metabólito.

Os dados dos ensaios de bioequivalência usualmente são provenientes de planejamentos “crossover” com períodos definidos previamente. O mais utilizado é o planejamento “crossover” 2x2, isto é, com duas amostras de tamanho mínimo 12, de acordo com as especificações da ANVISA, alocadas aleatoriamente para as formulações de referência (R) e teste (T) e cruzadas em dois períodos. O modelo “crossover” de bioequivalência pressupõe efeitos de seqüência e períodos não-significantes. Ao longo deste trabalho serão discutidos os principais conceitos referentes aos estudos de bioequivalência. No capítulo 2 encontra-se uma breve revisão de literatura sobre os estudos clínicos e os métodos bioanalíticos, como uma tentativa de introduzir o leitor ao assunto. No mesmo capítulo é apresentada a caracterização do planejamento experimental “crossover” de acordo com as recomendações da ANVISA. No capítulo 3 são abordadas as diferentes formas de análise, passando pela estruturação dos dados, as formas de análise, intervalos de confiança, testes de hipóteses, dimensionamento de amostra, procedimentos bootstrap, análise de medidas repetidas e as considerações adicionais, que incluem tópicos sobre outliers, análise não-paramétrica, uso de covariáveis, valores faltantes e bioequivalência individual e populacional. Já no capítulo 4 são aplicadas todas as técnicas descritas nos itens

anteriores a partir de dados hipotéticos disponíveis no Manual da ANVISA, onde todas as análises são descritas passo a passo. Por último, o capítulo 5 trata dos procedimentos computacionais de análise disponíveis no mercado.

Em anexo segue um glossário com os principais conceitos para que o leitor se familiarize com a linguagem técnica.

2 PLANEJAMENTO

Esse capítulo foi baseado no Manual de Boas Práticas em Biodisponibilidade e Bioequivalência da ANVISA.

2.1 Ensaios Clínicos

Ensaio Clínico é qualquer forma de experiência planejada que envolva pacientes e seja programada para elucidar qual o melhor tratamento para futuros pacientes com uma determinada condição clínica. A característica essencial de um ensaio clínico é que este utiliza resultados baseados numa amostra limitada de pacientes, de forma a inferir sobre qual o tratamento que deve ser aplicado à população de pacientes que necessitará de tratamento no futuro.

Segundo o manual, os estudos clínicos utilizando drogas em seres humanos podem ser divididos convencionalmente em 4 fases:

– Fase I

É o primeiro estudo em seres humanos. Envolve cerca de 20 a 50 indivíduos, sendo geralmente voluntários sadios, de acordo com a classe da droga a ser avaliada. São estudos de farmacologia clínica, nos quais se busca avaliar suas características de segurança e do perfil farmacocinético.

– Fase II

Estes estudos constituem a primeira administração do medicamento a pacientes, envolvendo cerca de 50 a 300 indivíduos. Têm como objetivo estudar o potencial terapêutico e os efeitos colaterais do medicamento, além de estabelecer as suas relações dose-resposta para empregá-las em ensaios terapêuticos mais definitivos.

– Fase III

São estudos terapêuticos multicêntricos, envolvendo no mínimo 250 indivíduos (usualmente este número chega a 3000/4000 pacientes), avaliando a eficácia da droga, sua segurança e comparando-as com placebo ou drogas já disponíveis no mercado com a mesma finalidade terapêutica. Exploram-se nesta fase o tipo e perfil das reações adversas mais freqüentes.

– Fase IV

Geralmente, são estudos de vigilância pós-comercialização, para estabelecer o valor terapêutico, o surgimento de novas reações adversas e/ou confirmação da frequência de surgimento das já conhecidas, e as estratégias de tratamento. São realizados com base nas características com que foi autorizado o medicamento e/ou especialidade medicinal.

Assim, entre os estudos que buscam identificar parâmetros iniciais de eficácia e segurança, estão os estudos de biodisponibilidade, quando os dados de absorção, distribuição, metabolização e eliminação, bem como efeitos adversos ainda não são totalmente conhecidos. Nestes estudos, uma monitorização médica rigorosa durante toda fase de investigação da droga faz-se necessária. Já os estudos de bioequivalência (biodisponibilidade relativa), por outro lado, apresentam, como principal propósito, obter evidências de que uma formulação teste não é diferente, do ponto de vista farmacocinético, de uma dada formulação referência. Estes estudos são realizados geralmente como base para solicitação de registro de um medicamento genérico. São conduzidos habitualmente em voluntários sadios, ou seja, um estudo não terapêutico, onde as características farmacocinéticas e farmacodinâmicas da droga estudada já são conhecidas, apresentando assim um risco menor que estudos iniciais de biodisponibilidade.

2.2 Método Bioanalítico

2.2.1 Introdução

A realização de uma pesquisa bibliográfica é a primeira etapa para a busca do método bioanalítico, que são procedimentos padrões para a realização de planejamentos experimentais. São divididos em três etapas: clínica, analítica e estatística. A etapa clínica envolve a seleção e o recrutamento dos voluntários para o estudo; a etapa analítica é responsável pela determinação da concentração dos medicamentos testados nas amostras de sangue coletadas, onde se utilizam métodos validados e as normas das Boas Práticas de Laboratório. E a etapa estatística analisa a existência ou não de diferenças significativas entre o comportamento das drogas testadas, concluindo assim sobre a bioequivalência. Uma vez existindo o método, ele deverá ser testado quanto a sua reprodutibilidade. Na inexistência de um método bioanalítico para um determinado fármaco, o centro analítico deve desenvolver um método que responda satisfatoriamente ao estudo desejado.

A realização prévia das etapas necessárias no desenvolvimento do método analítico para os estudos de bioequivalência assegura ao centro analítico e ao seu contratante que os serviços contratados serão realizados no tempo previsto e com a confiabilidade necessária dos resultados, os quais serão avaliados para fins de registro do medicamento em estudo. Neste contexto pode-se afirmar que contratado e contratante não perderão tempo e nem recursos financeiros adicionais se por acaso os estudos realizados forem rejeitados no seu término em função da inadequabilidade do método utilizado e das condições de armazenamento não determinadas.

2.2.2 Pré-validação

Os estudos de bioequivalência empregam a utilização de voluntários humanos, em número relativamente elevado, que não podem ser envolvidos no estudo sem a certeza de que os seus sacrifícios resultem num benefício para a sociedade. A certeza desses resultados para os voluntários só pode ser garantida se o método analítico para realização do estudo tiver sido desenvolvido previamente de forma a assegurar que os fluídos biológicos obtidos dos voluntários serão devidamente analisados.

No desenvolvimento de um método é necessário verificar toda a metodologia de preparação da amostra, a qual envolve os processos de extração, separação, purificação, identificação e quantificação do fármaco na matriz biológica. Para tanto, alguns estudos preliminares de validação devem ser efetuados visando à determinação das seguintes características: (1) exatidão, precisão e recuperação; (2) linearidade e limites de quantificação e (3) seletividade.

2.2.2.1 Exatidão, Precisão e Recuperação

A exatidão de um método analítico descreve a proximidade dos resultados médios obtidos pelo método em relação ao valor verdadeiro (concentração) do fármaco. A exatidão é determinada pela análise de amostras em replicata (repetição), contendo quantidades conhecidas do analito. A exatidão deve ser medida usando um mínimo de 5 determinações por concentração.

A precisão de um método analítico descreve a proximidade entre as diferentes medidas individuais de um fármaco. A precisão em termos de repetibilidade é obtida quando as amostras são preparadas nas mesmas condições de manuseio: lote, operador, equipamento, instrumento e

ocasião. Quando as amostras são preparadas com diferentes condições de manuseio, a precisão medida é a reprodutibilidade. No desenvolvimento do método, a precisão deve ser medida usando um mínimo de três concentrações com cinco repetições.

A recuperação de um fármaco de uma matriz biológica é a quantidade do analito obtida após o processo de separação. A quantificação é comparada com a concentração nominal de um padrão adicionada à matriz biológica antes do processo de separação. A recuperação indica a eficiência de todos os processos envolvidos no método analítico e deve ser tratada dentro de um limite de variabilidade. A recuperação de um fármaco não necessita ser de 100%, mas deve ser consistente, precisa e reprodutível. Quanto mais próxima de 100% for a recuperação, melhor o método analítico.

2.2.2.2 Linearidade e Limites de Quantificação

O método analítico em desenvolvimento deve atender aos requisitos de sensibilidade e detectabilidade planejados para o fármaco em função das suas concentrações plasmáticas na curva farmacocinética. Dessa forma é necessário determinar a relação entre a concentração do fármaco em função da resposta do detector, visando a determinação da faixa de concentração, onde a resposta é linear, e os respectivos limites de quantificação.

2.2.2.3 Especificidade e Seletividade

Especificidade é a busca individualizada do sinal de uma espécie única do analito. Seletividade é a capacidade de um método analítico de diferenciar e quantificar o analito na presença de outros compostos na amostra, denominados de interferentes.

Os interferentes em fluidos biológicos podem ser originários das fontes endo/exógenas. Os interferentes endógenos são os metabólitos e/ou precursores; produtos de degradação do fármaco, co-administração de fármacos, vitaminas e/ou seus metabólitos, produtos de interação de fármaco, componente biológico e outras substâncias que ocorrem naturalmente em fluidos biológicos, isto é, hormônios, proteínas, lipídios, substâncias dietéticas, etc. Os interferentes exógenos são as impurezas dos reagentes usados, substâncias liberadas pelos recipientes em uso ou resultantes de lavagem inadequadas de vidrarias, equipamentos e instrumentos.

Os níveis de interferentes devem ser avaliados antes de iniciar o desenvolvimento do método. Uma vez que o método esteja no estágio de validação, o teste de seletividade deve assegurar níveis e a reprodutibilidade dos interferentes e o seu impacto na precisão e exatidão do método em relação ao limite de quantificação.

Uma vez que os interferentes tenham sido definidos e minimizados durante o desenvolvimento do método nos estudos preliminares com a matriz biológica, deve-se, então, planejar um estudo para melhorar a avaliação estatística da extensão dos interferentes das matrizes biológicas nos diferentes voluntários. Na prática, interferência no estudo de pré-dose das amostras é um dos problemas mais comumente observados com os métodos analíticos validados, uma vez que eles são colocados em rotina. Isto surge da avaliação de um número insuficiente de voluntários durante a validação. Pela conferência de Washington recomenda-se uma avaliação de seis voluntários.

2.2.3 Estabilidade dos Fármacos em Fluidos Biológicos

O desenvolvimento do método é a etapa mais importante nos estudos analíticos da bioequivalência, visto que o comportamento dos fármacos nas matrizes biológicas depende dos níveis de interferentes que interagem com a molécula ativa. Um outro aspecto a ressaltar é a mudança dos constituintes da matriz biológica submetida a processo de estocagem levando em consideração o tempo e a temperatura. Assim sendo, produtos de degradação, complexação, oxidação, metabólitos e outras substâncias alteram a resposta de um método se o mesmo não for suficientemente seletivo para os estudos com matrizes biológicas frescas e envelhecidas.

Os estudos prévios de estabilidade em matrizes biológicas são essenciais para fornecer os parâmetros de adequabilidade do método em amostras envelhecidas e condições de armazenamento das amostras para os estudos analíticos da bioequivalência. Tais estudos permitem a confiabilidade necessária ao método, pois os resultados a serem fornecidos não correrão risco de serem rejeitados por erros de metodologia, quando a amostra é submetida a processo de envelhecimento. Por outro lado o armazenamento da amostra em uma dada temperatura por um tempo determinado deve ser baseado em dados científicos que possam afirmar que naquelas condições de estocagem o fármaco não sofrerá alteração.

2.2.4 Validação do Método

a. Validação Total

A validação total é importante no desenvolvimento e implementação de um método quando aplicado pela primeira vez ou quando for utilizado para quantificar um novo analito nesta mesma condição analítica.

b. Validação Parcial

As validações parciais são modificações do método já validado. Uma validação parcial pode compreender desde uma pequena determinação de precisão/exatidão a até quase uma validação total. Algumas mudanças típicas no método se enquadram nesta categoria, mas não são limitadas a: transferências de método bioanalítico entre laboratórios ou analistas; mudança na metodologia analítica (ex. mudança no sistema de detecção); mudança no anticoagulante na coleta do fluido biológico; mudança no processamento das amostras; mudança relevante na faixa de concentração; mudanças nos instrumentos e/ou software; volume de amostra limitado (ex. estudo pediátrico); demonstração seletiva de um analito na presença de medicações concomitantes.

c. Ferramentas para validação

As ferramentas usadas para a validação dos métodos bioanalíticos se dividem em dois grupos. O primeiro engloba as equações estatísticas, e dentre elas, pode-se citar a estimativa da variância, a estimativa do desvio padrão e o coeficiente de variação. E o segundo representa os testes de rejeição Teste Q e Teste de Grubbs.

2.3 Planejamento Experimental

De acordo com a Resolução nº 898, de 29 de maio de 2003, da ANVISA, fica determinado que a etapa de planejamento e as demais etapas estatísticas para Ensaio de Bioequivalência e Biodisponibilidade devem seguir as normas do “Guia para Planejamento e Realização da Etapa Estatística de Estudos de Biodisponibilidade Relativa/Bioequivalência”. O objetivo deste guia é fornecer algumas recomendações gerais para análise estatística nos estudos em questão.

Um dos critérios para escolher um delineamento apropriado é verificar se o delineamento selecionado pode identificar e isolar a variabilidade inter-individual (entre os sujeitos) na análise de dados. Qualquer delineamento que remova essa variação da comparação entre formulações pode ser apropriado.

O planejamento experimental mais utilizado nos ensaios de biodisponibilidade relativa e bioequivalência é o planejamento cruzado ou de reversão simples ("crossover design"). Mais detalhes sobre o modelo "crossover" em Jones & Kenward (1990).

2.4 Caracterização do Planejamento Experimental

Segundo a Agência Nacional de Vigilância Sanitária (ANVISA), em estudos de bioequivalência, tem-se interesse na quantidade e velocidade de absorção do fármaco analisado no sistema circulatório. Esses estudos são aleatorizados, comparativos e cruzados.

Um delineamento básico para um estudo de bioequivalência deve considerar importantes aspectos, como questões científicas a serem respondidas; natureza do material de referência e a forma farmacêutica a ser testada; disponibilidade de métodos analíticos e considerações do benefício do teste em seres humanos. Além dessas quatro questões há outras mais específicas, como o delineamento experimental, a aleatorização, o cronograma de coleta de dados, o período de eliminação e o número de voluntários.

2.4.1 Delineamento experimental

O planejamento adequado do experimento deve ter como objetivo minimizar a variabilidade que pode advir de várias fontes: variabilidade inter-individual; variabilidade intra-individual; efeito dos períodos, que pode ser causado por ação residual de tratamentos precedentes; erro experimental e variabilidade associada a tratamentos diferentes, como administração de produtos ou dosagens diferentes.

2.4.2 Aleatorização

Inferências estatísticas válidas são normalmente baseadas nas suposições de que os erros do modelo empregado são variáveis aleatórias independentemente distribuídas, o que pode ser assegurado através da aleatorização, que é muito importante para futuras comparações.

2.4.3 Período de Eliminação (“washout”) e Efeitos Residuais (“Carry-Over Effects”)

A presença de efeitos residuais (“carry-over effects”) tem um grande impacto na inferência estatística de bioequivalência entre formulações. Uma das soluções para esses casos pode ser o aumento do número de seqüências (períodos) no experimento.

O período de eliminação (“washout”) é definido como um intervalo de tempo suficientemente grande entre dois períodos de administração para que o efeito residual de uma formulação administrada num período seja eliminado até o próximo período.

Se um fármaco tem uma meia vida longa ou se o intervalo entre os períodos de tratamento é muito curto, o efeito do mesmo pode persistir depois do fim de período de eliminação, e esse efeito é conhecido como efeito residual. Neste caso, é necessário distinguir a diferença entre o efeito do fármaco e os efeitos residuais. O efeito do fármaco é aquele observado durante o período no qual ele é administrado, enquanto que o efeito residual é o efeito da droga que persiste depois do fim do período de dosagem.

2.4.4 Número de voluntários

O número de voluntários sadios deverá sempre assegurar poder estatístico suficiente para garantir a confiabilidade dos resultados. O poder mínimo padrão é de 80%.

2.4.5 Seleção do Delineamento Experimental

Selecionar um delineamento experimental apropriado ao planejar um estudo de biodisponibilidade relativa ou bioequivalência é uma questão importante. A resposta dessa questão depende de vários fatores, tais como o número de formulações a serem comparadas; as características do fármaco e sua biodisponibilidade; o objetivo do estudo; a variabilidade inter e

intra-individuais; a duração do estudo e número de períodos empregados; o custo de adição de um voluntário relativo à adição de um período e a taxa de desistência, também chamada de dropout.

2.5 O Planejamento “Crossover”

O delineamento “crossover” é uma variante do planejamento quadrado latino no qual cada coluna (bloco) recebe mais de uma formulação de um mesmo fármaco em períodos diferentes. Uma coluna (bloco) pode ser um sujeito ou um grupo de sujeitos. Os sujeitos em cada coluna ou bloco recebem uma seqüência diferente de formulações.

As vantagens em se utilizar o planejamento “crossover” para estudos de bioequivalência são: cada sujeito serve como seu próprio controle, o que permite uma comparação do sujeito com ele mesmo, para as diferentes formulações; remoção da variabilidade inter-individual da comparação entre formulações, o que torna o teste de diferença de tratamentos em geral mais poderoso e, com uma aleatorização apropriada de sujeitos para a seqüência de administração das formulações, o planejamento produz as melhores estimativas não viciadas para diferença (ou razão) entre formulações.

Um aspecto importante também a considerar é o balanceamento do delineamento. Para um delineamento ser dito balanceado, ele deve satisfazer as seguintes condições: cada medicamento é aplicado somente uma vez em cada voluntário (sujeito); em cada período, o número de voluntários que recebem cada medicamento tem que ser igual e o número de voluntários que recebem o medicamento i em algum período seguido pelo medicamento j no período seguinte é o mesmo para todo i diferente de j . O balanceamento é importante para a utilização das fórmulas propostas.

2.6 Tipos de Delineamentos “Crossover”

2.6.1. Delineamento “Crossover” para Dois Medicamentos

Nesse caso testam-se dois medicamentos: teste (T) e referência (R). Todos delineamentos apresentados neste trabalho são balanceados. Os delineamentos crossover para este tipo de configuração são os seguintes:

a) Delineamento crossover 2x2

É um delineamento convencional não replicado com duas formulações, dois períodos e duas seqüências. A sua configuração é:

Seqüência	Período	
	1	2
1	R	T
2	T	R

Cada sujeito é aleatoriamente alocado para a seqüência RT ou TR em dois períodos. Isto é, sujeitos alocados na seqüência RT (TR) recebem a formulação R (T) no primeiro período de administração e a formulação T (R) no segundo. Os períodos são separados por um período de eliminação adequado para evitar a ocorrência de efeito residual.

b) Delineamento crossover replicado

Este delineamento é recomendado para estudos de bioequivalência de produtos com fármacos de alta variabilidade (coeficiente de variação intra-individual superior a 30%), incluindo aqueles que são de liberação imediata (efeito do composto químico é liberado logo após ingestão), liberação modificada (quando ao usarmos de tecnologia farmacêutica, incorporamos junto ao fármaco novos efeitos, fazendo com que ocorra uma modificação na absorção e liberação do mesmo) e outros produtos de administração oral.

Para este delineamento os mesmos lotes das formulações teste e referência devem ser usados para a administração replicada. Os períodos devem ser suficientemente espaçados para garantir a inexistência do efeito residual.

Os delineamentos crossover replicados mais comumente usados para comparar duas formulações são:

(i) Delineamento com quatro seqüências e dois períodos (delineamento de Balaam):

Seqüência	Período	
	1	2
1	R	T
2	T	R
3	R	T
4	T	R

(ii) Delineamento com duas seqüências e quatro períodos:

Seqüência	Período			
	1	2	3	4
1	T	R	R	T
2	R	T	T	R

(iii) Delineamento com quatro seqüências e quatro períodos:

Seqüência	Período			
	1	2	3	4
1	T	T	R	R
2	R	R	T	T
3	T	R	R	T
4	R	T	T	R

(iv) Delineamento com duas seqüências e três períodos:

Seqüência	Período		
	1	2	3
1	T	R	R
2	R	T	T

ou

Seqüência	Período		
	1	2	3
1	T	R	T
2	R	T	R

Um número maior de voluntários é recomendado para o delineamento de três períodos, comparado com o delineamento de quatro períodos, para poder alcançar o mesmo poder estatístico para o teste.

2.6.2 Delineamento Crossover para Três Medicamentos

Esse delineamento é conhecido como delineamento de Williams, que possibilita testar os medicamentos teste 1 (T1), teste 2 (T2) e referência (R).

Para comparar três formulações de um fármaco, existem três possíveis pares de comparações: formulação 1 versus formulação 2, formulação 1 versus formulação 3 e formulação 2 versus formulação 3. Quando o número de formulações a serem comparadas é grande, mais seqüências e conseqüentemente mais indivíduos serão necessários, o que pode ser inviável. Um delineamento de uso prático proposto por (Williams (1949) Apud ANVISA), possui propriedades de balanceamento e requer poucas seqüências e períodos, o que pode ser uma boa alternativa. Um delineamento de Williams é dado por:

Seqüência	Período		
	1	2	3
1	R	T2	T1
2	T1	R	T2
3	T2	T1	R
4	T1	T2	R
5	T2	R	T1
6	R	T1	T2

2.6.3 Delineamento Crossover para Quatro Medicamentos

Esse delineamento também é conhecido como delineamento de Williams, em que se testam quatro medicamentos, teste 1 (T1), teste 2 (T2), teste 3 (T3) e referência (R), cuja ilustração é:

Seqüência	Período			
	1	2	3	4
1	R	T3	T1	T2
2	T1	R	T2	T3
3	T2	T1	T3	R
4	T3	T2	R	T1

3 ANÁLISE DE ENSAIOS DE BIOEQUIVALÊNCIA

3.1 Os Dados

Esse capítulo segue a estruturação e recomendações da ANVISA em seu Manual de Boas Práticas em Biodisponibilidade e Bioequivalência.

Os níveis de concentração de um fármaco no organismo dependem em parte da via de administração, que pode ser classificada como intravascular ou extravascular. A administração intravascular se dá diretamente na corrente sangüínea por via intravenosa ou intra-arterial. A extravascular inclui os meios de administração oral, intramuscular, subcutâneo, transdérmico e outros. Quando um fármaco é administrado no organismo, o mesmo geralmente atravessa as fases de absorção, distribuição, metabolização, e finalmente, eliminação. Dessa forma, a biodisponibilidade é geralmente determinada pelas medidas farmacocinéticas, ou seja, aquelas características que se relacionam com a quantidade de fármaco absorvido e com a velocidade do processo de absorção. Essas medidas podem ser obtidas a partir de resultados da quantificação do fármaco em líquidos biológicos, como sangue ou urina, após administração extravascular com dose única.

3.1.1 Principais Medidas Farmacocinéticas

De acordo com a Resolução nº 478, de 19 de março de 2002, publicada no Diário Oficial de 20/03/2002, em Estudos de Bioequivalência para Medicamentos Genéricos, os parâmetros ou medidas farmacocinéticas (variáveis resposta, no contexto estatístico) deverão ser obtidas das curvas de concentração sangüínea do fármaco em função do tempo, e analisadas estatisticamente para determinação de bioequivalência. Essa resolução também estabelece três medidas fundamentais à determinação da bioequivalência entre medicamentos: a área sob a curva de concentração em função do tempo (ASC), a concentração máxima observada (C_{max}), e o tempo no qual essa concentração foi alcançada (T_{max}). A curva de concentração do fármaco é caracterizada pela quantificação de um determinado número de amostras biológicas, relativas a tempos de coleta previamente estabelecidos.

A primeira e mais importante medida avaliada é a área sob a curva de concentração plasmática do fármaco em função do tempo, freqüentemente utilizada para medir a extensão da

absorção, ou o montante total de droga absorvido pelo organismo, após administração de dose única de um medicamento. A determinação da bioequivalência entre dois medicamentos resulta da comparação das ASC's obtidas no experimento. Através de sua representação matemática

$$ASC = \frac{FD}{K_e Vd} \quad (1)$$

observa-se que esta medida é diretamente proporcional à quantidade de fármaco efetivamente absorvido e disponível para ser distribuído no organismo (FD). O termo $K_e Vd$, dado em (1), expressa o *clearance* total do fármaco (depuração), ou seja, sua velocidade de eliminação a partir do volume de distribuição, sendo que K_e corresponde à constante de velocidade de eliminação total do fármaco no organismo.

Dentre diversos métodos para a determinação da ASC do tempo zero até o tempo da última coleta (t_k), o mais utilizado é o método dos trapezóides usado por (Chow & Liu (2000) Apud ANVISA). Esse método consiste na soma das áreas dos trapézios determinados pelos tempos de coleta e respectivas concentrações. Sejam $C_0, C_1, C_2, \dots, C_k$, as concentrações obtidas em um experimento para os tempos de coleta $0, t_1, t_2, \dots, t_k$, respectivamente. A ASC de zero a t_k , denotada por ASC_{t_k} , é obtida da seguinte forma:

$$ASC_{t_k} = \sum_{t=1}^k \left(\frac{C_{t-1} + C_t}{2} \right) (t - t_{t-1}) \quad (2)$$

A área sob a curva de concentração em função do tempo (ASC) pode também ser extrapolada e calculada do tempo zero até o tempo relativo à completa eliminação do fármaco. Essa medida é citada na literatura como a área sob a curva do tempo zero a infinito. A porção adicional é expressa por uma relação entre a última concentração medida C_k e a constante de velocidade de eliminação do fármaco K_e . A constante de eliminação é calculada para cada voluntário como o coeficiente de inclinação da reta de regressão ajustada nos 4 a 6 últimos valores de concentração transformados em logaritmo de base 10, multiplicadas por uma

constante de transformação de base $-2,303$. A área sob a curva de zero a infinito é obtida da seguinte maneira:

$$ASC_{\infty} = ASC_{t_k} + \frac{C_k}{K_e} \quad (3)$$

A ocorrência de alguns valores faltantes e/ou valores inesperados na curva de concentração, geralmente não acarretará um grande impacto no cálculo de ASC. Por outro lado, se esses valores forem relativos aos últimos pontos da curva, como por exemplo t_k , a estimativa da ASC pode vir a ter um viés. De acordo com a legislação vigente no Brasil, a área sob a curva do tempo zero ao tempo t_k , deve ser igual ou superior a 80% da área sob a curva de zero a infinito. O pico de concentração máxima (Cmax), que representa a maior concentração do fármaco observada é diretamente proporcional ao montante total de droga absorvido pelo organismo.

O tempo no qual a concentração máxima foi alcançada (Tmax), é o tempo de coleta no qual foi observada a ocorrência de Cmax, e relaciona-se com a velocidade de absorção do fármaco. Os valores de Cmax e Tmax são obtidos diretamente sem interpolação dos dados. A determinação da bioequivalência entre dois medicamentos inclui também a comparação das medidas de Cmax e Tmax obtidas no experimento para cada uma das formulações (medicamentos teste e referência).

Adicionalmente também são obtidos a depuração (D), o volume aparente de distribuição (Vd) e a meia-vida de eliminação ($t_{1/2}$) do fármaco e/ou metabólito, muito embora para essas medidas farmacocinéticas não se tenha necessidade de tratamento estatístico.

O volume aparente de distribuição (Vd) é uma medida que correlaciona a quantidade de fármaco no organismo com a sua concentração (C) no sangue ou no plasma, dependendo do fluido biológico que se esteja trabalhando. O volume não se refere a um valor fisiológico identificável, mas simplesmente ao volume de líquido que seria necessário para conter todo o fármaco no organismo na mesma concentração presente no sangue ou no plasma. Dessa forma, o volume de distribuição de determinado fármaco pode variar em função da idade, sexo, doença e composição orgânica de cada pessoa. Nesse sentido, são extremamente importantes os critérios de inclusão do estudo para a seleção dos voluntários, na tentativa de se utilizar um grupo com características mais homogêneas possíveis.

Outra medida farmacocinética é a meia-vida ($t_{1/2}$) de um fármaco, que é o tempo gasto para que a concentração plasmática reduza-se à metade, sendo definida por

$$t_{1/2} = \frac{\log 2}{K_e} \quad (4)$$

Embora a meia-vida seja considerada um parâmetro fraco para a determinação da eliminação de um fármaco, ela proporciona uma boa indicação sobre o tempo gasto para que um fármaco seja removido do organismo, funcionando como um meio para se calcular os intervalos entre as doses de administração de um medicamento. No caso da bioequivalência, essa medida funciona como uma ferramenta essencial para a determinação do tempo de descanso necessário para a completa eliminação do medicamento ingerido no primeiro período do estudo. A legislação brasileira prevê que o intervalo entre os períodos do estudo deve ser de no mínimo sete vezes à meia-vida do fármaco estudado.

A depuração (D) ou *clearance* de um determinado indivíduo é a medida da capacidade do organismo em eliminar um fármaco, e é dada pelo produto do volume aparente de distribuição e da constante de eliminação:

$$D = V_d * K_e \quad (5).$$

3.1.2 Estudos Com Doses Múltiplas

No caso de estudos com doses múltiplas deve-se comprovar que o estado de equilíbrio foi alcançado após a administração dos medicamentos teste e de referência.

Para estudos que empregam doses múltiplas devem também ser determinados:

- 1) ASC_{t_k} calculado no intervalo de dose (t_k) no estado de equilíbrio;
- 2) C_{max} e T_{max} , obtidos sem interpolação de dados;
- 3) Concentração mínima do fármaco (C_{min}), determinada no final de cada intervalo de dose do estado de equilíbrio;
- 4) Concentração média do fármaco no estado de equilíbrio

$$C^* = ASC_{t_k} / t_k \quad (6)$$

5) Grau de flutuação no estado de equilíbrio

$$GF = (C_{max} - C_{min}) / C^* \times 100 \quad (7)$$

3.1.3 Outras Considerações Importantes

Ainda, segundo a Resolução nº 478 da ANVISA, os parâmetros ASC_{t_k} e C_{max} devem ser transformados em logaritmo natural ou logaritmo na base 10, pois a distribuição destes dados transformados se aproxima mais de uma distribuição normal do que os dados originais. A escolha da base do logaritmo deve ser consistente e especificada no relatório de estudo.

A limitação do tamanho da amostra utilizada num estudo típico de bioequivalência impede uma determinação confiável da distribuição do conjunto de dados. A utilização da transformação logarítmica justifica-se em relação ao tratamento estatístico dos dados e quanto aos aspectos farmacocinéticos.

Em geral, uma comparação preliminar de interesse num estudo de bioequivalência é a utilização da razão ao invés da diferença, entre as médias dos parâmetros farmacocinéticos (ASC e C_{max}) dos dados dos produtos teste e referência. Usando transformação logarítmica, o modelo empregado na análise de dados permite fazer inferências estatísticas sobre a diferença entre duas médias na escala logarítmica, as quais podem ser re-transformadas propiciando inferências estatísticas sobre a razão das duas médias na escala original. (Schuirmann (1989) Apud ANVISA).

De acordo com o manual da ANVISA, Westlake (1973, 1988) observou que um modelo multiplicativo é adequado para medidas farmacocinéticas (ASC e C_{max}) num estudo de bioequivalência. Assumindo que a eliminação do fármaco é de primeira ordem e ocorre somente a partir do compartimento central, a ASC , dada de acordo com a equação (1), é obtida após uma administração extravascular (oral). Portanto, o uso de ASC como uma medida da quantidade de medicamento absorvido envolve um termo multiplicativo ($CL = K_e V_d$), o qual pode ser considerado como uma função do voluntário. Esse modelo não é aditivo para os dados na escala original. A transformação logarítmica da ASC resulta num modelo aditivo:

$$\log ASC_y = \log F + \log D - \log V - \log K_e \quad (8)$$

Argumentos semelhantes foram apresentados no caso de C_{max} . Já para o T_{max} , a análise deverá ser feita pela diferença individual (teste – referência), construindo um intervalo com 95% de confiança, utilizando-se um procedimento não-paramétrico, devido à alta variabilidade dessa característica e pelo fato de sua distribuição ser desconhecida.

3.2 Métodos de Análise de Bioequivalência

3.2.1 Introdução

Este capítulo segue a estruturação do Manual de Boas Práticas em Biodisponibilidade e Bioequivalência da ANVISA. Outras fontes de referência importantes são Fleiss (1999), John (1971) e Jones & Kenward (1990).

A análise de bioequivalência entre duas formulações de referência (R) e teste (T) tem sido realizada com base nas medidas farmacocinéticas (variáveis respostas, na linguagem estatística) comentadas em 3.1.1. Os métodos de análise estatística conhecidos concentram-se em bioequivalência média, individual e populacional. O primeiro considera a análise sob o enfoque das médias das formulações R e T, enquanto que os dois últimos consideram a variabilidade existente nas medidas farmacocinéticas. O método da bioequivalência média é o mais comumente usado e será abordado neste capítulo.

Os métodos estatísticos de bioequivalência média surgiram no fim da década de 70 e início da década de 80, como consequência dos esforços do FDA de apontar as necessidades de métodos apropriados de avaliação estatística dos estudos de bioequivalência. Os principais desenvolvimentos foram a reformulação das hipóteses de bioequivalência, o método do intervalo de confiança e a abordagem bayesiana. Diversos estudos têm sido realizados para a análise de bioequivalência com estudos comparativos dos métodos.

Antes da comparação da bioequivalência e biodisponibilidade das formulações, é interessante realizar alguns testes preliminares para verificar a presença de efeito do período e de efeitos residuais, pois a ocorrência de efeitos residuais pode certamente aumentar a complexidade da análise estatística.

3.2.2 Modelo Estatístico

Num planejamento crossover, é importante remover os efeitos residuais da comparação já que o efeito de droga pode ser confundido com qualquer efeito residual. Entretanto, em algumas situações isso não é possível. Levando em conta os efeitos residuais, o seguinte modelo estatístico é normalmente considerado.

Seja Y_{ijk} a resposta (por exemplo, $ASCT_k$, Cmax) do i -ésimo indivíduo na k -ésima seqüência no j -ésimo período:

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{(j,k)} + R_{(j-1,k)} + \varepsilon_{ijk} \quad (9)$$

onde:

μ : média geral;

S_{ik} : efeito do i -ésimo sujeito na k -ésima seqüência, onde i (sujeito) = 1, 2, 3, ..., n_k e k (seqüência) = 1, 2, ..., g ;

P_j : efeito fixo do j -ésimo período, $j = 1, 2, \dots, p$ tal que $\sum_j P_j = 0$;

$F_{(j,k)}$: efeito fixo da formulação administrada na k -ésima seqüência e j -ésimo período tal que, $\sum F_{(j,k)} = 0$;

$R_{(j-1,k)}$: efeito residual fixo de primeira ordem da formulação administrada na k -ésima seqüência e $(j-1)$ -ésimo período, tal que $R_{(0,k)} = 0$ e $\sum R_{(j-1,k)} = 0$;

ε_{ijk} : erro aleatório (dentro de sujeito).

Assume-se que $\{S_{ik}\}$ e $\{\varepsilon_{ijk}\}$ são independentes, identicamente distribuídos (i.i.d) com média zero e variâncias σ_s^2 e σ_t^2 respectivamente, onde $t=1,2, \dots, L$ (número de formulações a serem comparadas). Assume-se também que $\{S_{ik}\}$ e $\{\varepsilon_{ijk}\}$ são mutuamente independentes. As estimativas de σ_s^2 são geralmente usadas para explicar a variabilidade inter-individual (inter-

sujeitos), enquanto que as estimativas de σ_i^2 são usadas para descrever a variabilidade intra-individual (intra-sujeitos), para a t-ésima formulação.

Também pode-se incorporar ao modelo os efeitos residuais de segunda ordem e/ou efeitos da interação direta tratamento por período no modelo.

3.2.3 Inferência Estatística sobre Efeitos de um Delineamento Crossover 2x2

Com base no modelo apresentado na seção 3.2.2 equação (9), vemos que o delineamento crossover 2x2 é uma variação do mesmo, onde os indivíduos variam de $i = 1, 2, \dots, n_k$; os períodos de $j = 1, 2$ e a seqüência $k = 1, 2$. Nesse caso, a notação pode ser simplificada de modo que $F_{(j,k)} = F_R$ ou F_T e $R_{(j-1,k)} = R_R$ ou R_T .

Como a formulação administrada no primeiro período na primeira seqüência é de referência, então:

$F_{(j,k)} = F_R$ para $k = j$ (efeito fixo da formulação na k-ésima seqüência e no j-ésimo período é igual à formulação referência se $k = j$)

$F_{(j,k)} = F_T$ para $k \neq j$ (efeito fixo da formulação na k-ésima seqüência e no j-ésimo período é igual à formulação teste se $k \neq j$).

Para esse delineamento os efeitos residuais podem ocorrer somente no segundo período. Seja R_R o efeito residual da formulação referência do primeiro para o segundo período na seqüência 1. Portanto,

$$R_{(j-1,k)} = R_R \text{ se } k = 1 \text{ e } j = 2;$$

$$R_{(j-1,k)} = R_T \text{ se } k = 2 \text{ e } j = 2.$$

Para cada sujeito, os efeitos fixos em cada período e seqüência são dados por:

Quadro 1. Efeitos fixos para as seqüências e períodos.

Seqüência	Período 1	Período 2
1 (RT)	$\mu_{11} = \mu + P_1 + F_R$	$\mu_{21} = \mu + P_2 + F_T + R_R$
2 (TR)	$\mu_{12} = \mu + P_1 + F_T$	$\mu_{22} = \mu + P_2 + F_R + R_T$

Onde $\mu_{jk} = E(Y_{ijk})$, $P_1 + P_2 = 0$, $F_R + F_T = 0$ e $R_R + R_T = 0$.

Para a comparação de biodisponibilidade entre formulações, é desejável estimar e separar os efeitos residuais do efeito da droga. Mas na prática é normal assumir que não existe efeito de período e que não existem efeitos residuais. Isso se deve ao fato de que um estudo bem conduzido pode eliminar o possível efeito de período, além de que um período de eliminação (washout) suficientemente espaçado pode assegurar a ausência de efeitos residuais. Entretanto, em vários casos, os efeitos de período e/ou efeitos residuais ainda podem permanecer.

3.2.3.1 Efeitos Residuais

Considere $U_{ik} = Y_{i1k} + Y_{i2k}$ como a soma da medida farmacocinética em estudo nos períodos 1 e 2 na k-ésima seqüência, onde o índice i denota os sujeitos e varia de 1 a n_k e k representa a seqüência ($k = 1, 2$). A média da soma na k-ésima seqüência será:

$$\bar{U}_{.k} = \frac{1}{n_k} \sum U_{ik}, k = 1, 2 \quad (10)$$

onde $\bar{U}_{.1}$ e $\bar{U}_{.2}$ são variáveis aleatórias independentes provenientes de populações normais com variâncias iguais.

Seja $R = R_T - R_R$. Sob a restrição de que $R_T + R_R = 0$, os efeitos residuais são iguais para duas formulações, isto é, $R = 0$ se e somente se $R_T = R_R = 0$. Portanto, um teste para ausência de efeitos residuais é equivalente a um teste de igualdade de efeitos residuais. Em outras palavras, a existência de efeito residual significa que existem efeitos residuais diferentes nas seqüências de tratamentos. Contudo, mas a inexistência do efeito residual não implica necessariamente que tais efeitos sejam nulos, mas que se existirem, têm a mesma intensidade em ambas seqüências de tratamentos.

Quando há inexistência de efeitos residuais, os efeitos de droga (isto é, $F = F_T - F_R$) podem ser estimados com base nos dados de ambos períodos. Mas não há estimadores não-viciados para os efeitos de droga caso os efeitos residuais estejam presentes. Portanto, é importante examinar a existência desses efeitos.

Os efeitos residuais podem ser determinados testando-se as hipóteses:

$$H_0 : R = 0 \quad (\text{ou } R_T = R_R)$$

$$H_1 : R \neq 0 \quad (\text{ou } R_T \neq R_R)$$

A rejeição da hipótese de nulidade nos leva à conclusão da presença de efeitos residuais entre formulações.

Define-se $\hat{R} = \bar{U}_{.2} - \bar{U}_{.1}$ como uma variável normalmente distribuída com média R e variância $Var\hat{R} = \sigma_u^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$ estimada por:

$$V\hat{ar}(\hat{R}) = \hat{\sigma}_u^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \quad (11)$$

onde

$$\hat{\sigma}_u^2 = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^2 \sum_{i=1}^{n_k} (U_{ik} - \bar{U}_{.k})^2. \quad (12)$$

Sob H_0 $R = 0$, a estatística

$$T_r = \frac{\hat{R}}{\sqrt{(\hat{\sigma}_u^2) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (13)$$

tem uma distribuição t de Student com $n_1 + n_2 - 2$ graus de liberdade.

Rejeita-se a hipótese nula H_0 ao nível α se $|T_r| > t_{\alpha/2}(n_1 + n_2 - 2)$, onde $t_{\alpha/2}(n_1 + n_2 - 2)$ é o valor crítico da distribuição t de Student, ao nível α de significância e $n_1 + n_2 - 2$ graus de liberdade.

Como a estatística T_r envolve a estimativa de $\sigma_u^2 = 2(\sigma_s^2 + \sigma_i^2)$ que inclui as variabilidades inter e intra individuais, ela pode ter pouco poder quando a variabilidade inter-individual (inter-sujeitos) é relativamente maior que a intra-individual. Para aumentar o poder do teste, (Grizzle (1965) Apud ANVISA) sugeriu testar H_0 usando $\alpha = 0,10$.

3.2.3.2 O efeito da Droga

Consideremos a metade da diferença das medidas em estudo entre os períodos 1 e 2 na k -ésima seqüência:

$$d_{ik} = \frac{1}{2}(Y_{i2k} - Y_{i1k}), i = 1, \dots, n_k; k = 1, 2 \quad (14)$$

e a média destas diferenças na seqüência k

$$\bar{d}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} d_{ik}; k = 1, 2. \quad (15)$$

Denota-se o efeito da droga por $F = F_T - F_R$. Se $R_T \neq R_R$, então não existem estimadores não-viciados para F baseado nos dados de ambos os períodos. Por outro lado, se $R_T = R_R$, então

$$\hat{F} = \bar{d}_{.1} - \bar{d}_{.2} = \bar{Y}_T - \bar{Y}_R \quad (16)$$

onde

$$\bar{Y}_R = \frac{1}{2}(\bar{Y}_{.11} + \bar{Y}_{.22}) \quad (17)$$

$$\bar{Y}_T = \frac{1}{2}(\bar{Y}_{.21} + \bar{Y}_{.12}) \quad (18)$$

\bar{Y}_T e \bar{Y}_R são as chamadas médias de mínimos quadrados para as formulações T e R, respectivamente. Pode-se mostrar que \hat{F} é normalmente distribuído com média F e variância

$Var\hat{F} = \sigma_d^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$, onde a estimativa de σ_d^2 é dada por:

$$\hat{\sigma}_d^2 = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^2 \sum_{i=1}^{n_k} (d_{ik} - \bar{d}_k)^2 \quad (19)$$

A presença do efeito da droga pode ser examinada testando-se as hipóteses

$$H_0 : F_R = F_T$$

$$H_1 : F_R \neq F_T$$

A través da estatística

$$T_d = \frac{\hat{F}}{\sqrt{(\hat{\sigma}_d^2) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (20)$$

testa-se a hipótese de nulidade. Rejeita-se H_0 ao nível α de significância se $|T_r| > t_{\alpha/2}(n_1 + n_2 - 2)$, onde $t_{\alpha/2}(n_1 + n_2 - 2)$ é o valor crítico da distribuição t de Student, ao nível α de significância e $n_1 + n_2 - 2$ graus de liberdade.

Na presença de efeitos residuais, um estimador não-viciado para F pode ser obtido usando-se somente os dados do primeiro período. Neste caso, um intervalo de $(1 - \alpha)100\%$ de confiança para F é dado por:

$$\hat{F}^* \pm t(\alpha/2; n_1 + n_2 - 2) \hat{\sigma}_d \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad (21)$$

onde

$$\hat{F}^* = \bar{Y}_{.12} - \bar{Y}_{.11} \quad (22)$$

e

$$\hat{\sigma}_d^2 = \frac{1}{n_1 + n_2 - 2} \sum_{k=1}^2 \sum_{i=1}^{n_k} (Y_{ilk} - \bar{Y}_{.1k})^2. \quad (23)$$

Assim, quando existe efeito residual, a hipótese nula de inexistência de efeito da droga é rejeitada se $|T_r| > t_{\alpha/2}(n_1 + n_2 - 2)$, onde

$$T_{d/r} = \frac{\hat{F}^*}{\hat{\sigma}_d \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (24)$$

é o valor da estatística de teste e t é o valor crítico da distribuição t de Student com $n_1 + n_2 - 2$ graus de liberdade e nível de significância α .

Na prática, quando existem efeitos residuais, os dados referentes ao primeiro período podem ser analisados para avaliação da bioequivalência entre as formulações. Entretanto, podem vir à tona várias conseqüências, como o pouco poder para detecção de uma diferença clinicamente significativa devido ao aumento da variabilidade e também o fato de que ao desconsiderar-se a informação do segundo período, elimina-se o benefício de um delineamento crossover, que remove a variabilidade inter-individual da comparação entre formulações.

3.2.3.3 O Efeito de Período

Um estimador não-viciado para o efeito de período P pode ser obtido como:

$$\hat{P} = \frac{1}{2} \left[(\bar{Y}_{21} - \bar{Y}_{11}) - (\bar{Y}_{12} - \bar{Y}_{22}) \right]. \quad (25)$$

As hipóteses a serem testadas são:

$$H_0 : P_1 = P_2$$

$$H_1 : P_1 \neq P_2.$$

Rejeita-se H_0 ao nível de significância α se $|T_r| > t_{\alpha/2}(n_1 + n_2 - 2)$ onde

$$T_p = \frac{\hat{P}}{\hat{\sigma}_d \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (26)$$

é a estatística de teste e t é o valor crítico da distribuição t de Student com $n_1 + n_2 - 2$ graus de liberdade e nível de significância α .

3.2.4 Análise de Variância

O objetivo da análise de variância é estudar a variabilidade nos dados observados, particionando a soma de quadrados total das observações em componentes relacionados aos efeitos fixos e aleatórios do modelo.

Com base no modelo apresentado em 3.2.2, pode-se notar que os efeitos de sujeito, período, droga e efeito residual de primeira ordem, que são fixos, serão testados contra o erro aleatório (resíduo), conforme mostra a Tabela 1. Pode-se perceber também que o efeito de sujeitos ou indivíduos é uma composição de dois efeitos: seqüência e sujeitos (seqüência), que correspondem, respectivamente ao efeito residual com 1 grau de liberdade e ao erro entre os sujeitos dentro de cada seqüência, de acordo com o apresentado no Manual da ANVISA. A componente resíduo (erro) representa a variabilidade dentro dos sujeitos. A Tabela 1 apresenta a Anova para o modelo crossover generalizado para L número de tratamentos a serem testados nos p períodos de tempo, e considerando n o número de repetições do quadrado latino, que equivale à metade do tamanho amostral total. Nessa tabela, a fonte de variação efeito residual fixo ($R_{(j-1,k)}$) não se encontra isolada, mas com o seu efeito incorporado ao efeito de sujeitos.

Tabela 1. Anova para o delineamento “crossover”.

Fonte de Variação	Graus de liberdade	SQ	QM	E(QM)	F	Hipóteses
Tratamento (droga)	L-1	$\sum_{j=1}^p \frac{Y_{j\cdot}^2}{np} - \frac{Y_{\dots}^2}{N}$	$\frac{SQ_{trat}}{p-1}$	$\sigma^2 + np \sum_{j=1}^p F_j^2 / (p-1)$	$\frac{QM_{trat}}{QME}$	$H_0 : F_j = 0$ para todo j
Linhas (períodos)	p-1	$\sum_{i=1}^p \sum_{l=1}^n \frac{Y_{i\cdot l}^2}{p} - \frac{Y_{\dots}^2}{N}$	$\frac{SQ_{per}}{p-1}$	$\sigma^2 + np \sum_{i=1}^p P_i^2 / (p-1)$	$\frac{QM_{per}}{QME}$	$H_0 : P_j = 0$ para todo j
Colunas (sujeitos)	np-1	$\sum_{k=1}^p \frac{Y_{\cdot k}^2}{np} - \frac{Y_{\dots}^2}{N}$	$\frac{SQ_{suj}}{np-1}$	$\sigma^2 + p \sum_{k=1}^p S_k^2 / (np-1)$	$\frac{QM_{suj}}{QME}$	$H_0 : S_k = 0$ para algum k
Resíduo	(p-1)(np-2)	Por subtração	$\frac{SQE}{(p-1)(np-2)}$	σ^2		
Total	$np^2 - 1$	$\sum_i \sum_j \sum_k \sum_l Y_{ijkl}^2 - \frac{Y_{\dots}^2}{N}$				

A tabela de análise de variância para um delineamento crossover 2x2, adotada pela ANVISA, foi criada por (Hills & Armitage (1979) Apud ANVISA). De acordo com o manual da ANVISA, (Grizzle (1965) Apud ANVISA) foi o primeiro a apresentar a tabela de Anova para o modelo crossover 2x2, mas seus resultados foram apenas corretos para o caso especial em que $n_1 = n_2$. Anos depois, apresentou uma correção, mas, segundo (Grieve (1982) Apud ANVISA), ainda apresentava inconsistências.

Conforme mostra a Tabela 2, pode-se perceber que o delineamento “crossover” 2x2 é um caso particular de parcela subdividida (split-plot), onde os sujeitos formam a parcela principal, os períodos de tempo, como medidas repetidas, são consideradas as parcelas secundárias (subdivisões), e os efeitos são testados com diferentes denominadores (erro(A) e erro(B)). A Anova está particionada em duas fontes de variação: entre indivíduos e dentro de indivíduos. A variabilidade entre os indivíduos engloba a variação da seqüência, e o efeito dos indivíduos hierárquico à seqüência. O efeito de seqüência será testado com os resíduos entre os indivíduos, e caso não apresente significância estatística, pode ter seu efeito incorporado ao efeito de indivíduos, formando uma única fonte de variação entre os indivíduos. A outra fonte de variação (dentro de indivíduos) apresenta os efeitos de droga, período e resíduos dentro de indivíduos, onde as somas de quadrado para efeito de droga são ajustadas para períodos e as somas de quadrado para efeito de período são ajustadas para tratamentos (drogas). Os efeitos de droga e período serão testados com os resíduos dentro de indivíduos. A Tabela 2 apresenta a Anova para um planejamento “crossover” 2x2.

Tabela 2. Anova para planejamento “crossover” 2x2.

Fonte de Variação	GL	SQ	QM	F
Entre Indivíduos				
Efeito Residual (seqüência)	1	$SQ_{residual}$	$SQ_{residual}$	$F_r = \frac{QM_{residual}}{QM_{inter}}$
Resíduo (entre indivíduos)	n-2	SQ_{inter}	$\frac{SQ_{inter}}{n-2}$	
Dentro de Indivíduos				
Droga (tratamento)	1	SQ_{trat}	SQ_{trat}	$F_{trat} = \frac{QM_{trat}}{QM_{int ra}}$
Período	1	$SQ_{periodo}$	$SQ_{periodo}$	$F_p = \frac{QM_{periodo}}{QM_{int ra}}$
Resíduo (dentro de indivíduos)	n-2	$SQ_{int ra}$	$\frac{SQ_{int ra}}{n-2}$	
Total	2n-1	SQ_{total}		

Seja \bar{Y}_{\dots} a média geral das observações. Então

$$SQ_{total} = \sum_{i=1}^{n_1} \sum_{j=1}^2 \sum_{k=1}^2 (Y_{ijk} - \bar{Y}_{\dots})^2 = \sum_{i=1}^{n_1} \sum_{j=1}^2 \sum_{k=1}^2 (Y_{ijk} - \bar{Y}_{i,k})^2 + 2 \sum_{k=1}^2 \sum_{i=1}^{n_1} (\bar{Y}_{i,k} - \bar{Y}_{\dots})^2 \quad (27)$$

onde

$$\bar{Y}_{i,k} = \frac{1}{2} \sum_{j=1}^2 Y_{ijk} \quad (28)$$

$$SQ_{residual} = \frac{2n_1 n_2}{n_1 + n_2} \left\{ \frac{1}{2} [(\bar{Y}_{.12} + \bar{Y}_{.22}) - (\bar{Y}_{.11} + \bar{Y}_{.21})] \right\}^2 \quad (29)$$

$$SQ_{inter} = \sum_{k=1}^2 \sum_{i=1}^{n_1} \frac{Y_{i,k}^2}{2} - \sum_{k=1}^2 \frac{Y_{..k}^2}{2n_k} \quad (30)$$

$$SQ_{trat} = \frac{2n_1n_2}{n_1 + n_2} \left\{ \frac{1}{2} [(\bar{Y}_{.21} - \bar{Y}_{.11}) - (\bar{Y}_{.22} - \bar{Y}_{.12})] \right\}^2 \quad (31)$$

$$SQ_{per} = \frac{2n_1n_2}{n_1 + n_2} \left\{ \frac{1}{2} [(\bar{Y}_{.21} - \bar{Y}_{.11}) - (\bar{Y}_{.12} - \bar{Y}_{.22})] \right\}^2 \quad (32)$$

e

$$SQ_{intera} = \sum_{i=1}^{n_y} \sum_{j=1}^2 \sum_{k=1}^2 Y_{ijk}^2 - \sum_{k=1}^2 \sum_{i=1}^{n_y} \frac{Y_{i,k}^2}{2} - \sum_{k=1}^2 \sum_{j=1}^2 \frac{Y_{.jk}^2}{n_k} + \sum_{k=1}^2 \frac{Y_{.k}^2}{2n_k} \quad (33).$$

Em ambas tabelas pode-se afirmar que o principal interesse é a significância do efeito de droga. Os demais efeitos podem ser considerados secundários. Se necessário, também pode ser adicionada ao modelo a interação simples entre os tratamentos (droga) e período.

3.3 Métodos Estatísticos para Bioequivalência Média

3.3.1 Introdução

Este capítulo tem como principais referências o Manual da ANVISA, Jones & Kenward (1990).

O objetivo dos estudos de biodisponibilidade é verificar se a taxa e a extensão da absorção do fármaco no sistema circulatório referentes às duas formulações referência e teste (R e T) são equivalentes. Sabe-se de estudos anteriores que os perfis de biodisponibilidade de duas formulações podem variar para um mesmo indivíduo e por isto trabalha-se com uma avaliação em termos médios. Se ao considerar-se as médias populacionais de alguma das medidas farmacocinéticas de biodisponibilidade, as duas formulações são consideradas bioequivalentes se a diferença ou a razão entre as duas médias diferir por menos que os limites especificados. Normalmente usa-se a regra de 20% para cima ou para baixo como limites de tolerância. Este método de verificar a bioequivalência é denominado bioequivalência média. Existem outros métodos de verificar a bioequivalência (individual e populacional), mas neste capítulo serão descritos apenas os métodos de bioequivalência média, que são baseados em testes de hipóteses e intervalos de confiança para a diferença ou a razão das médias. Em algumas situações há a

necessidade de se trabalhar com uma transformação das medidas farmacocinéticas, sendo a logarítmica a mais usada, e por esse motivo, as análises serão separadas em dois grupos: com e sem transformação.

3.3.2 Intervalos para Bioequivalência

Nesta seção serão apresentados os limites de bioequivalência baseados na diferença e na razão das médias para os casos de medidas farmacocinéticas originais ou transformadas.

3.3.2.1 Medidas Farmocinéticas sem Transformação

Para essa situação são utilizadas as médias aritméticas das medidas farmacocinéticas μ_R e μ_T para definir os intervalos de bioequivalência entre R e T.

O intervalo de bioequivalência para a diferença das médias é expresso por $\theta_{\text{inf}} < \mu_T - \mu_R < \theta_{\text{sup}}$. Utilizando a regra dos limites de tolerância $\pm 20\%$, os limites inferior e superior de bioequivalência são:

$$\theta_{\text{inf}} = -0,20\mu_R \text{ e } \theta_{\text{sup}} = 0,20\mu_R \quad (34)$$

e então o intervalo de bioequivalência fica determinado como:

$$-0,20\mu_R < \mu_T - \mu_R < 0,20\mu_R \quad (35)$$

Na realidade esses parâmetros devem ser estimados, pois dependem do parâmetro populacional, e a escolha natural acaba se tornando $\hat{\theta}_{\text{inf}} = -0,20\bar{y}_R$ e $\hat{\theta}_{\text{sup}} = 0,20\bar{y}_R$, onde \bar{y}_R é a média amostral da formulação de referência (R). Assim, o intervalo de bioequivalência real é o seguinte

$$-0,20\bar{y}_R < \mu_T - \mu_R < 0,20\bar{y}_R \quad (36)$$

De forma semelhante, o intervalo de bioequivalência para a razão das médias é expresso por $\delta_{\text{inf}} < \mu_T / \mu_R < \delta_{\text{sup}}$. De (35), o intervalo de bioequivalência para a diferença das médias aritméticas para as duas formulações é $-0,20\mu_R < \mu_T - \mu_R < 0,20\mu_R$, então dividindo-se por μ_R , obtêm-se $-0,20 < (\mu_T - \mu_R) / \mu_R < 0,20$ ou $-0,20 < \mu_T / \mu_R - 1 < 0,20$, ou seja, $0,80 < \mu_T / \mu_R < 1,20$ (37). Assim os limites inferior e superior de bioequivalência são, respectivamente: $\delta_{\text{inf}} = 0,80$ e $\delta_{\text{sup}} = 1,20$.

3.3.2.2 Medidas Farmocinéticas com Transformação Logarítmica

Denotando por $\tilde{\mu}$ a média geométrica da medida farmacocinética na escala original X e μ a média aritmética da medida transformada de X na base neperiana ($Y = \ln X$), pode-se mostrar que $\mu = \ln \tilde{\mu}$. E então, na escala original a média geométrica é $\tilde{\mu} = \exp(\mu)$. Essas relações aparecem de forma organizada na Tabela 3.

Tabela 3. Relações entre as médias.

Escala	Média	Relação entre $\tilde{\mu}$ e μ
Original: X	Geométrica: $\tilde{\mu}$	$\tilde{\mu} = \exp(\mu)$
Logarítmica: Y = lnX	Aritmética: μ	$\mu = \ln \tilde{\mu}$

Através dessas transformações pode-se mostrar que o intervalo de bioequivalência para a razão das médias definido por $\delta_{\text{inf}} < \tilde{\mu}_T / \tilde{\mu}_R < \delta_{\text{sup}}$ tem como limites inferior e superior $\delta_{\text{inf}} = 0,80$ e $\delta_{\text{sup}} = 1,25$. Então o intervalo fica expresso por

$$0,80 < \frac{\tilde{\mu}_T}{\tilde{\mu}_R} < 1,25. \quad (38)$$

O uso dos limites 0,80 e 1,25 para intervalo de bioequivalência da razão das médias na escala original é justificado pelo fato de haver uma correspondência a um intervalo de bioequivalência simétrico para as diferenças na escala transformada.

O intervalo de bioequivalência para a razão das médias, ao utilizar-se uma transformação logarítmica é dado por $\ln 0,80 < \ln(\tilde{\mu}_T / \tilde{\mu}_R) < \ln(1,25)$, ou ainda

$-0,2231 < \ln \tilde{\mu}_T - \ln \tilde{\mu}_R < 0,2231$. Logo o intervalo de bioequivalência das diferenças das médias aritméticas na escala transformada é expresso por:

$$-0,2231 < \mu_T - \mu_R < 0,2231 \quad (39)$$

onde os limites inferior e superior são respectivamente $-0,2331$ e $+0,2331$.

A Tabela 4 apresenta um resumo dos limites dos intervalos de bioequivalência para as quatro situações apresentadas.

Tabela 4. Intervalos de bioequivalência baseados na média aritmética ou geométrica para a medida farmacocinética original ou transformada.

Escala	Diferença	Razão
	$\theta_{\text{inf}} < \mu_T - \mu_R < \theta_{\text{sup}}$	$\delta_{\text{inf}} < \mu_T / \mu_R < \delta_{\text{sup}}$
Original	$-0,20\bar{y}_R < \mu_T - \mu_R < 0,20\bar{y}_R$	$80\% < \mu_T / \mu_R < 120\%$
Logarítmica	$-0,2231 < \mu_T - \mu_R < 0,2231$	$80\% < \frac{\tilde{\mu}_T}{\tilde{\mu}_R} < 125\%$

Note que o caso da diferença na escala original é o único em que os limites de bioequivalência dependem dos dados, os demais já são fixos.

3.3.3 Avaliação da Bioequivalência Média

A comparação de médias (diferença ou razão) pode ser realizada através de métodos paramétricos ou não-paramétricos. Para a utilização de métodos paramétricos, supõe-se que a variável de interesse tenha um determinado tipo de distribuição. Em estudos de bioequivalência, geralmente a medida farmacocinética com ou sem transformação pode ser descrita pela distribuição normal. Já para a utilização de métodos não-paramétricos, não há exigências sobre a distribuição da variável estudada e por esta razão, são também chamados de métodos de distribuição livre. Portanto, boas opções aplicáveis para situações em que não se conhece a distribuição ou quando a variável não segue a distribuição normal.

A avaliação da bioequivalência pode ser feita através de intervalos de confiança (IC) ou teste de hipóteses. Existem vários métodos paramétricos e não-paramétricos para a construção de intervalos de confiança e de teste de hipóteses. A seguir serão apresentados somente os métodos paramétricos e não-paramétricos padrões para cada tipo de inferência estatística.

3.3.3.1 Intervalo de Confiança

Se um intervalo com $100(1-\alpha)\%$ de confiança para a diferença ($\mu_T - \mu_R$) ou para a razão (μ_T / μ_R) está dentro dos limites aceitáveis recomendados pela agência reguladora (ANVISA), isto é, nos intervalos $[\theta_{inf}; \theta_{sup}]$ ou $[\delta_{inf}; \delta_{sup}]$, respectivamente, então a conclusão é de que há bioequivalência, caso contrário, conclui-se pela não existência de bioequivalência.

Se a regra de $\pm 20\%$ é adotada para medidas farmacocinéticas sem transformação, o nível de significância é usualmente escolhido como 5% e o coeficiente de confiança como 95% para testar as diferenças e razões de médias. Para as medidas farmacocinéticas com transformação logarítmica, usando a regra de 80/125, o nível de significância a ser escolhido também é usualmente de 5% e a confiança de 95%.

Para a decisão de bioequivalência média, a regra adotada é de que o intervalo de confiança esteja contido no intervalo de bioequivalência. A chance de que o intervalo com $100(1-\alpha)\%$ de confiança esteja dentro dos limites de equivalência não é necessariamente maior ou igual a $1-\alpha$. Se o coeficiente de variação (CV) for alto, o método de intervalo de confiança para decisão de bioequivalência média pode não ter o nível desejado requerido pela agência reguladora, nesse caso a ANVISA (CV de 14% a 22%).

A seguir serão apresentados os métodos padrões paramétricos e não-paramétricos, para a construção do intervalo de confiança para a diferença e a razão de médias. O método considerado será o do intervalo com menor amplitude.

(A) Método paramétrico de intervalo com menor amplitude (“shortest”)

A1. Medida farmacocinética sem transformação

Para a situação de diferença de médias, e supondo normalidade da distribuição da medida farmacocinética a ser analisada, o intervalo de confiança para a diferença das médias $\mu_T - \mu_R$, com nível de confiança de $100(1 - \alpha)\%$ é $[LD_{inf}; LD_{sup}]$, sendo os limites calculados como:

$$\left[(\bar{y}_T - \bar{y}_R) - t_{\alpha/2}(n_1 + n_2 - 2) * s_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}; (\bar{y}_T - \bar{y}_R) + t_{\alpha/2}(n_1 + n_2 - 2) * s_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] \quad (40)$$

onde

\bar{y}_T : média de mínimos quadrados da formulação teste;

\bar{y}_R : média de mínimos quadrados da formulação referência;

$t_{\alpha/2}(n_1 + n_2 - 2)$: percentil de confiança $(1 - \alpha)$ da distribuição t de Student, com $n_1 + n_2 - 2$ graus de liberdade;

$s_d = (QM_{int\ ra} / 2)^{1/2}$: desvio-padrão ponderado da diferença de período de ambas as seqüências, onde $QM_{int\ ra}$ é a variância intra-individual;

n_1 : tamanho de amostra da primeira seqüência;

n_2 : tamanho de amostra da segunda seqüência.

E para o caso de razão de médias μ_T / μ_R , o intervalo com nível de confiança $100(1 - \alpha)\%$ é calculado a partir dos limites do intervalo de confiança para a diferença das médias $[LD_{inf}; LD_{sup}]$. Assumindo que μ_R possa ser substituído por \bar{y}_R , o intervalo de confiança expresso em percentuais $[LR_{inf}; LR_{sup}]$ é dado por

$$\left[\left(\frac{LD_{inf}}{\bar{y}_R} + 1 \right) 100\%; \left(\frac{LD_{sup}}{\bar{y}_R} + 1 \right) 100\% \right] \quad (41)$$

A2. Medida farmacocinética com transformação logarítmica

Supondo que na escala logarítmica a medida farmacocinética siga a distribuição normal, o intervalo com $100(1-\alpha)\%$ de confiança para a diferença das médias aritméticas dos logaritmos das medidas farmacocinéticas de T e R, $\mu_T - \mu_R$ é obtido da mesma forma que o caso sem transformação. Basta substituir em (42) os valores das estatísticas calculadas na escala logarítmica. O intervalo será denotado por:

$$[LD_{inf}^{\sim}; LD_{sup}^{\sim}] \quad (42)$$

O intervalo com $100(1-\alpha)\%$ de confiança para a razão das médias μ_T / μ_R é calculado a partir dos limites do intervalo de confiança para a diferença das médias $(LD_{inf}; LD_{sup})$, dado por (38) e seus limites $[LR_{inf}^{\sim}; LR_{sup}^{\sim}]$ são expressos por:

$$[\exp\{LD_{inf}\}; \exp\{LD_{sup}\}]. \quad (43)$$

Resumindo, a regra para decidir sobre bioequivalência com intervalos de confiança paramétricos é a seguinte:

Conclui-se por bioequivalência se

- $(LD_{inf}; LD_{sup})$ ou $(LD_{inf}^{\sim}; LD_{sup}^{\sim})$ está contido no intervalo $(\theta_{inf}; \theta_{sup})$ ou
- $(LR_{inf}; LR_{sup})$ ou $(LR_{inf}^{\sim}; LR_{sup}^{\sim})$ está contido em $(\delta_{inf}; \delta_{sup})$,

onde os limites aparecem na Tabela 5.

Tabela 5. Condições de bioequivalência para a diferença e a razão das médias com medida farmacocinética com e sem transformação.

Escala	Diferença	Razão
Original	$(LD_{inf}; LD_{sup}) \subset (-0,20\bar{y}_R; 0,20\bar{y}_R)$	$(LR_{inf}; LR_{sup}) \subset (80\%; 120\%)$
Logarítmica	$(LD_{inf}^{\sim}; LD_{sup}^{\sim}) \subset (-0,2231; 0,2231)$	$(LR_{inf}^{\sim}; LR_{sup}^{\sim}) \subset (80\%; 125\%)$

(B) Método Não-Paramétrico baseado no Estimador de Hodges-Lehmann

O método não-paramétrico para construção do intervalo com $100(1 - \alpha)\%$ de confiança para a diferença de médias baseado no Estimador de Hodges-Lehmann, segue os seguintes passos:

1. Calcular todos os possíveis pares de diferenças das diferenças de períodos entre as seqüências 1 e 2. Isto é, para os índices $i = 1, 2, \dots, n_1$ e $j = 1, 2, \dots, n_2$, calcular $D_{i,j} = d_{i1} - d_{j2}$, que são estimativas de $\theta = \mu_T - \mu_R$ para todos os pares das diferenças;
2. Ordenar as $n_1 n_2$ diferenças: $D(1) < D(2) < \dots < D(n_1 n_2)$;
3. Calcular a mediana de $\{D(i), i = 1, 2, \dots, n_1 n_2\}$ que é o estimador pontual de $\theta = \mu_T - \mu_R$, conhecido como Estimador de Hodges-Lehmann.
4. Identificar os valores de $\{D(i), i = 1, 2, \dots, n_1 n_2\}$ correspondentes aos percentis de ordem $\alpha/2$ e $1 - \alpha/2$, denotadas por $D(w(\alpha/2))$ e $D(w(1 - \alpha/2))$, que serão os limites inferior e superior do intervalo de confiança deste método.

O intervalo de confiança para esse método é definido como:

$$[LD_{\text{inf}} = D(w(\alpha/2)); LD_{\text{sup}} = D(w(1 - \alpha/2) + 1)] \quad (44)$$

A decisão de bioequivalência é verificada se $D(w(\alpha/2)) > \theta_{\text{inf}}$ e $D(w(1 - \alpha/2)) < \theta_{\text{sup}}$.

3.3.3.2 Teste de Hipóteses

Para a verificação de bioequivalência entre as formulações R e T, o Manual da ANVISA propõe como hipóteses nula e alternativa, respectivamente H_0 contra H_1 :

H_0 : Não Bioequivalência entre as formulações T e R

H_1 : Bioequivalência entre as formulações T e R

Em estudos de bioequivalência é usual o erro do tipo I (alfa) ou nível de significância igual a 5% e confiança de 95%.

Schuirmann, em 1981, introduziu uma notação própria para a formulação das hipóteses, que acabou servindo de base para o Manual da ANVISA (Schuirmann (1981) Apud ANVISA). O

autor sugeriu o uso de hipóteses de intervalo de bioequivalência média, que podem ser formuladas como dois conjuntos de hipóteses unilaterais:

$$H_{01} : \mu_T - \mu_R \leq \theta_{\text{inf}} \quad \text{versus} \quad H_{11} : \mu_T - \mu_R > \theta_{\text{inf}}$$

$$H_{02} : \mu_T - \mu_R \geq \theta_{\text{sup}} \quad \text{versus} \quad H_{12} : \mu_T - \mu_R < \theta_{\text{sup}}$$

Quando a medida farmacocinética é analisada sem transformação, θ_{inf} e θ_{sup} são escolhidos como 20% de μ_R . Quando a medida farmacocinética é analisada com transformação logarítmica, θ_{inf} e θ_{sup} assumem valores -0,2231 e 0,2231, respectivamente.

As hipóteses também podem ser expressas em termos de razões, como segue:

$$H_{01} : \mu_T / \mu_R \leq \delta_{\text{inf}} \quad \text{versus} \quad H_{11} : \mu_T / \mu_R > \delta_{\text{inf}}$$

$$H_{02} : \mu_T / \mu_R \geq \delta_{\text{sup}} \quad \text{versus} \quad H_{12} : \mu_T / \mu_R < \delta_{\text{sup}}$$

Os valores de δ_{inf} e δ_{sup} quando a medida farmacocinética é analisada sem transformação são respectivamente 80% e 120%, segundo a regra de $\pm 20\%$, e são 80% e 125% quando se utilizam transformações logarítmicas. Ao concluir-se que $\theta_{\text{inf}} < \mu_T - \mu_R$, rejeita-se a primeira hipótese nula; e se $\mu_T - \mu_R < \theta_{\text{sup}}$ é rejeitada a segunda hipótese nula. Rejeitando-se ambas as hipóteses nulas, pode-se concluir que $\theta_{\text{inf}} < \mu_T - \mu_R < \theta_{\text{sup}}$ e, portanto, que as médias das formulações teste e referência são bioequivalentes em média. Raciocínio análogo é feito para as hipóteses de intervalo para razão das médias. As rejeições das hipóteses nulas implicam na conclusão de bioequivalência.

Os métodos de testes de hipóteses de intervalo podem ser paramétricos, quando se supõe normalidade da medida farmacocinética em estudo, ou não-paramétricos, quando a suposição de normalidade é violada. Embora existam vários testes para bioequivalência, neste trabalho serão abordados apenas os métodos paramétrico padrão e não-paramétrico padrão.

(A) Método Paramétrico - Dois testes unilaterais t

Sob suposição de normalidade, os dois conjuntos de hipóteses unilaterais podem ser testados através de testes t unilaterais. Conclui-se que T e R são bioequivalentes se:

$$T = \frac{(\bar{y}_T - \bar{y}_R) - \theta}{s_d \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} > |t_\alpha(n_1 + n_2 - 2)| \quad (45)$$

onde $t_\alpha(n_1 + n_2 - 2)$ é o valor da distribuição t com $n_1 + n_2 - 2$ graus de liberdade e nível de significância α .

O procedimento dos dois testes t unilaterais é operacionalmente equivalente ao intervalo de confiança clássico, isto é, se o intervalo de confiança clássico de $100(1 - \alpha)\%$ para a diferença das médias está dentro dos limites, então, ambas as hipóteses nulas conjuntas são também rejeitadas ao nível de significância α pelo procedimento dos dois testes t unilaterais.

(B) Método Não-Paramétrico – Dois testes unilaterais de Wilcoxon

As hipóteses conjuntas formuladas para a diferença das médias das formulações teste e referência são utilizadas no teste não-paramétrico de Wilcoxon, cujos cálculos seguem os seguintes passos:

1. Calcular as diferenças de períodos d_{ik} , $i = 1, 2, \dots, n_k$, $k = \text{seqüência} = 1, 2$;
2. Obter os estimadores de $\theta_h^* = (\mu_T - \mu_R) - \theta_h$ como funções lineares de d_{ik} , $i = 1, 2, \dots, n_k$, $k = 1, 2$: $b_{hik} = d_{ik} - \theta_h$ para indivíduos da seqüência 1 (RT) ou $b_{hik} = d_{ik}$ para indivíduos da seqüência 2 (TR). O quadro 2 mostra os valores das funções para $h = \text{inf, sup}$ para o indivíduo i na seqüência k :

Quadro 2. Valores das funções para os testes unilaterais de Wilcoxon.

RT	d_{i1}	$b_{inf,i,1} = d_{i1} - \theta_{inf}$	$b_{sup,i,1} = d_{i1} - \theta_{sup}$
TR	d_{i2}	$b_{inf,i,2} = d_{i2}$	$b_{sup,i,2} = d_{i2}$

3. Ordenar $b_{inf,i,1}, i = 1, 2, \dots, n_1$ e $b_{inf,i,2}, i = 1, 2, \dots, n_2$ numa amostra combinada e assinalar postos $R(b_{inf,i,k})$ aos valores ordenados.
4. Ordenar de $b_{sup,i,1}, i = 1, 2, \dots, n_1$ e $b_{sup,i,2}, i = 1, 2, \dots, n_2$ numa amostra combinada e assinalar postos $R(b_{sup,i,k})$ aos valores ordenados. Somar os postos $R(b_{inf,i,k})$ dos sujeitos da seqüência 1, obtendo R_{inf} e somar os postos $R(b_{sup,i,k})$ dos sujeitos da seqüência 1, obtendo R_{sup} .

O teste de Wilcoxon para as hipóteses de intervalo é baseado em R_{inf} e R_{sup} . Conclui-se pela bioequivalência média quando:

$$W_{inf} = R_{inf} - n_1(n_1 + 1)/2 > w(1 - \alpha) \quad (46)$$

e

$$W_{sup} = R_{sup} - n_1(n_1 + 1)/2 < w(\alpha) \quad (47)$$

onde $w(1 - \alpha) = n_1 n_2 - w(\alpha)$ e $w(\alpha)$ são os percentis das distribuições da estatística de Wilcoxon.

Quando $n_1 + n_2$ é grande e a razão $n_1/n_2 \cong 0,5$, a aproximação pela distribuição normal pode ser utilizada. Neste caso, a bioequivalência média é verificada quando:

$$Z_{inf} = [R_{inf} - 0,5 * n_1(n_1 + n_2 + 1)]/[0,0833 * n_1 * n_2(n_1 + n_2 + 1)]^{1/2} > Z_\alpha \quad (48)$$

e

$$Z_{sup} = [R_{sup} - 0,5 * n_1(n_1 + n_2 + 1)]/[0,0833 * n_1 * n_2(n_1 + n_2 + 1)]^{1/2} > -Z_\alpha \quad (49)$$

onde Z_α é o valor da normal padronizada, ao nível de significância α .

3.3.4 Poder dos Testes Unilaterais

O poder de um teste estatístico para verificar a bioequivalência entre duas formulações R e T é definido como a probabilidade de rejeitar a hipótese nula (não bioequivalência) quando a hipótese nula é falsa, ou seja, quando as formulações R e T são bioequivalentes de fato. O poder de um teste está ligado a duas importantes considerações: a primeira, para a fase de planejamento do estudo de bioequivalência, que é o cálculo do tamanho de amostra, e a segunda, para a comparação entre dois testes de hipóteses, por exemplo, teste paramétrico e não-paramétrico. Em planejamentos estatísticos normalmente deseja-se um poder mínimo de 80%. Para o cálculo do poder, em estudos de bioequivalência, o Manual da ANVISA apresenta um algoritmo, onde são necessários os valores de $\theta = \mu_T - \mu_R$, $\Delta = -\theta_L = \theta_U$ e σ_d^2 ou coeficiente de variação (CV %).

O poder empírico dos dois testes unilaterais t pode ser obtido através do seguinte procedimento:

Etapa 1: Gerar amostras de tamanho $n_1 + n_2$ de acordo com o modelo, com valores pré-estabelecidos de μ_R , μ_T , σ_d^2 e σ_s^2 .

Etapa 2: Para um dado Δ , calcular T_{inf} e T_{sup} .

Etapa 3: Repetir etapas 1 e 2 B vezes.

Etapa 4: O poder empírico é a proporção entre as b amostras, tais que $T_{inf} > t_{\alpha}(n_1 + n_2 - 2)$ e $T_{sup} < -t_{\alpha}(n_1 + n_2 - 2)$, onde $t_{\alpha}(n_1 + n_2 - 2)$ é o valor da distribuição t de Student com $n_1 + n_2 - 2$ graus de liberdade.

Para o caso não-paramétrico, o poder empírico dos dois testes unilaterais de Wilcoxon pode ser obtido de forma semelhante, através do seguinte procedimento:

Etapa 1: Gerar amostras de tamanho $n_1 + n_2$ de acordo com o modelo, com valores pré-estabelecidos de μ_R , μ_T , σ_d^2 e σ_s^2 .

Etapa 2: Para um dado Δ , calcular R_{inf} , R_{sup} , W_{inf} e W_{sup} .

Etapa 3: Repetir etapas 1 e 2 B vezes.

Etapa 4: O poder empírico é a proporção entre as b amostras, tais que $W_{inf} > Z_{\alpha}$ e $W_{sup} < -Z_{\alpha}$, onde z é o valor da distribuição normal padrão com nível de significância α .

Maiores informações sobre o poder dos testes em ensaios de bioequivalência podem ser encontradas em (Chow & Liu (2000) e Schuirmann (1987) Apud ANVISA), como consta no manual da ANVISA, onde é apresentada a função poder dos testes unilaterais t. O cálculo da função poder pode ser realizado por integração numérica ou através de simulação (poder empírico). Os autores também realizam um estudo de simulação para comparar os testes t unilaterais e de Wilcoxon utilizando os poderes empíricos dos dois testes para vários valores de $\theta = \Delta = 20\%$ e do coeficiente de variação (CV). Desses estudos concluiu-se que o poder do teste paramétrico é maior que o poder do teste não-paramétrico, embora as diferenças sejam inferiores a 20% em todos os casos considerados.

3.3.5 Dimensionamento de Amostra

Em estudos de bioequivalência há inúmeras questões a respeito do dimensionamento de tamanho de amostra, e visando resolver tais questões, a ANVISA recomenda realizar um pré-estudo do poder baseado na estimativa da variabilidade intra-individual a partir de estudos anteriores disponíveis. O tamanho de amostra apropriado é escolhido de forma a se garantir um poder desejado para a avaliação da bioequivalência dentro de limites considerados clinicamente importantes.

Teoricamente o tamanho de amostra necessário para realizar um estudo de bioequivalência necessita de algumas informações relativas ao fármaco em estudo, tais como a variabilidade do fármaco nos indivíduos, medida pelo coeficiente de variação CV (%) e o tamanho mínimo do efeito que se quer detectar (por exemplo, magnitude da diferença entre as médias das duas formulações). Estas informações definiriam o desenho a ser utilizado, e em consequência, o cálculo do tamanho de amostra requerido. Entretanto, a obtenção destas informações a priori teoricamente é bastante difícil, fazendo com que na prática seja realizado um estudo piloto.

A Resolução RE n° 478 de 19 de março de 2002 sugere $n = 24$ como o tamanho mínimo da amostra para estudos crossover 2x2 de bioequivalência média. Utilizando a fórmula aproximada de cálculo de tamanho de amostra de (Chow & Liu (1992) Apud ANVISA), nota-se que o tamanho de amostra 24 estaria dentro de uma faixa percentual de CV de 14% a 22% para $\Delta = 0\%$ ou 5% e poder = 80% ou 90%.

A representatividade da população na amostra, ou validade, é assegurada pelo planejamento científico da pesquisa. O processo de amostragem começa com a determinação do tamanho e com a coleta adequada.

Os elementos essenciais para o dimensionamento de amostras são: fixar um nível de confiança α , o tamanho mínimo do efeito que se quer detectar Δ e dentro do contexto de testes de hipóteses, acrescenta-se o poder estatístico do teste $1-\beta$. As fontes desses dados são estudos anteriores realizados em populações e em condições semelhantes. Em procedimentos mais simples, determina-se o tamanho de amostra adequado baseando-se apenas em uma variável de destaque no estudo, o que não garante que seja suficiente para as outras variáveis também de interesse. Como os estudos de bioequivalência são baseados em várias medidas farmacocinéticas, idealmente os cálculos devem ser feitos para cada uma delas.

Chow & Liu (2000) Apud ANVISA fornecem uma fórmula aproximada para calcular o tamanho da amostra baseado na função de poder do teste por hipótese de intervalo. Como o poder é simétrico em relação a zero, considera-se apenas o caso em que $\theta > 0$. Para um nível de confiança de $1-\alpha$, um poder de $1-\beta$, estimativa $\hat{\sigma}_d^2$ e limite de bioequivalência pré-fixado Δ quando $0 < \theta = \theta_0 < \Delta$, o tamanho da amostra é dado por:

$$n_1(\theta_0) \geq 2 [t_{(\alpha, 2n-2)} + t_{(\beta, 2n-2)}]^2 \left(\frac{\hat{\sigma}_d}{\Delta - \theta_0} \right)^2 \quad (50)$$

onde

n : tamanho de amostra;

$\Delta = -\theta_{\text{inf}} = \theta_{\text{sup}}$: limites de bioequivalência;

α e β : probabilidades dos erros do tipo I e II, respectivamente;

$t_{1-\alpha/2, v}$ e $t_{1-\alpha, v}$: percentis da distribuição t de Student com v graus de liberdade correspondente ao nível de confiança $1-\alpha$, respectivamente para hipóteses bilateral e unilateral;

$t_{1-\beta, v}$: designa o percentil da distribuição t de Student com v graus de liberdade correspondente a um poder $1-\beta$;

$\hat{\sigma}_d = QM_{int ra} / \sqrt{2}$: estimativa do desvio-padrão ponderado da diferença de período de ambas as seqüências.

Se a regra de $\pm 20\%$ for usada com $\Delta = 0,2 \mu_R$, torna-se

$$n_I(\theta_0) \geq [t_{(\alpha, 2n-2)} + t_{(\beta, 2n-2)}]^2 \left(\frac{CV}{20 - \theta'_0} \right)^2 \quad (51)$$

onde

$$\theta'_0 = 100 * \theta_0 / \mu_R. \quad (52)$$

3.4 Procedimentos Bootstrap

3.4.1 Introdução

Este capítulo foi escrito seguindo Albanese (1990) e Manly (1997).

O bootstrap é um procedimento geral de reamostragem introduzido por Efron (1979) para estimar a distribuição das estatísticas baseadas em observações independentes. A palavra 'bootstrap' ou 'bootstrapping' foi usada por Efron para seguir o espírito de (Tukey (1958) Apud Albanese (1990)), que usou o termo 'jackknifing' para descrever seus estudos posteriores sobre métodos de reamostragem ("resampling").

A essência dos métodos bootstrap é a idéia de que na ausência de alguma informação sobre a população, a distribuição dos valores encontrados na amostra aleatória de tamanho n da população seja o melhor guia para a distribuição populacional. Em outras palavras, a população infinita que consiste nos n valores amostrais observados, cada um com probabilidade $1/n$, é usada para modelar a população real desconhecida. A amostragem é com reamostragem, que é a única diferença na prática entre bootstrap e aleatorização em muitas aplicações.

Um dos principais temas em pesquisa nas técnicas bootstrap, nos últimos quinze anos, têm sido o desenvolvimento dos métodos para calcular limites de confiança válidos para parâmetros populacionais. Mais recentemente, os testes de significância bootstrap têm recebido muita importância, tornando-se um dos principais pontos de pesquisa.

3.4.2 Técnica Bootstrap

O método bootstrap não-paramétrico ou empírico consiste no seguinte:

Considere X_1, X_2, \dots, X_p variáveis aleatórias independentes e identicamente distribuídas (iid) oriundas de uma população com função de distribuição F desconhecida, e cujo principal objetivo é fazer inferências sobre o vetor β de parâmetros da população.

Seja $\hat{\beta}(x_1, x_2, \dots, x_p)$ um estimador de β baseado no tamanho de amostra n e \hat{F} a sua distribuição empírica.

A técnica bootstrap aproxima a distribuição amostral de β sobre F da distribuição amostral de $\hat{\beta}$ sobre \hat{F} . Este procedimento é realizado com base no método de Monte Carlo, que consiste na realização dos seguintes passos:

(1) Construa \hat{F} .

(2) Retire a amostra bootstrap, $x^*_1, x^*_2, \dots, x^*_p$ iid com função distribuição acumulada \hat{F} e calcule

$$\hat{\beta}^* = \hat{\beta}(X^*_1, X^*_2, \dots, X^*_p) \quad (53)$$

(3) Repita independentemente por B vezes o passo 2 (para valores de B grande), obtendo $\hat{\beta}^*_b$, $b = 1, 2, \dots, B$. A função distribuição de $\hat{\beta}$ é aproximada por

$$\hat{F}_B(y) = \mathbf{P}(\hat{\beta}^*_b \leq y) / B \quad (54)$$

A estimativa bootstrap de β baseada nas B replicações é a média das estimativas $\hat{\beta}^*_b$, ou seja,

$$\hat{\beta}^* = \sum \hat{\beta}^*_b / B \quad (55)$$

e a matriz bootstrap de variância – covariância pode ser descrita como

$$\sum_B = (B - 1)^{-1} \sum (\hat{\beta}^*_b - \hat{\beta}^*)(\hat{\beta}^*_b - \hat{\beta}^*)^t \quad (56)$$

Quando o número de replicações $B \rightarrow \infty$, $\hat{\beta}^*$ será aproximado à estimativa bootstrap de β e \sum_B é correspondente à estimativa bootstrap da matriz de variância – covariância Σ .

A estimativa bootstrap do vício ou viés é a diferença entre a estimativa do parâmetro $\hat{\beta}$ e a estimativa bootstrap $\hat{\beta}^*$, ou seja,

$$\text{vicio} = \hat{\beta} - \hat{\beta}^* \quad (57)$$

O resultado básico da teoria bootstrap é que as distribuições empíricas das estimativas dos parâmetros obtidas por esse método são assintoticamente as mesmas que as das distribuições amostrais daqueles parâmetros na amostragem da população da qual a amostra original foi retirada.

Não há nada que diga que o método bootstrap deva ser realizado não parametricamente. Se há razões para acreditar que a verdadeira distribuição de F é normal, por exemplo, pode-se estimar F por suas estimativas paramétricas de máxima verossimilhança \hat{F} .

Efron (1979) também sugere que o método da expansão das séries de Taylor possa ser utilizado para obter a média e a variância aproximadas da distribuição bootstrap de $\hat{\beta}^*$, mostrando que ambas diferem apenas em detalhes do método jackknife padrão.

(Efron (1983) Apud Albanese (1990)) discute diferentes tipos de intervalo de confiança usando a técnica bootstrap, mostrando que é necessário ter no mínimo 1000 amostras para computar o BC (intervalo do percentil corrigido para viés) e os intervalos BC_α , enquanto que para o método mais simples, intervalo para percentil, 250 replicações podem dar resultados úteis.

O intervalo para percentil é obtido fazendo $\beta \in (\hat{F}_\alpha^{-1}; \hat{F}_{1-\alpha}^{-1})$ como uma aproximação para o intervalo central com $(1 - \alpha)100\%$ de confiança para β .

3.4.3 Intervalos de Confiança Bootstrap em Estudos de Bioequivalência

O método mais simples de construção de intervalos de confiança é usando uma aproximação para a distribuição normal com uma variância estimada pelo método não-paramétrico delta. Se I_j representa o valor de influência empírico para a observação Y_j , então a variância aproximada é dada por

$$v_L = n^{-2} \sum l_j^2. \quad (58)$$

Desse modo, o intervalo de confiança para o parâmetro θ de interesse é

$$t \pm v_L^{1/2} z(1-\alpha) \quad (59)$$

em que t é um estimador de θ . Se uma amostra bootstrap de tamanho pequeno é gerada, pode-se ter um viés de estimação. Para isto, pode-se calcular o viés b_R e a variância v_R , e o intervalo de confiança correspondente com coeficiente de confiança $(1-\alpha)$ é dado por:

$$t - b_R \pm v_L^{1/2} z(1-\alpha). \quad (60)$$

Desse modo, tem-se o intervalo de confiança calculado com uma correção de viés.

Quando se considera um método bootstrap básico, a única mudança que se faz para obter um intervalo de confiança não-paramétrico é considerar a função de distribuição empírica \hat{F} . Este método pode ou não fornecer uma melhoria no método anterior, através da obtenção do Q-Q-plot normal dos valores bootstrap t^* . Desse modo, um intervalo de confiança bootstrap “studentized” pode ser aplicado. Considerando a variância não-paramétrica v_L calculada pelo método delta, então os limites de confiança são dados por:

$$t \pm v_L^{1/2} z^*((R+1)(1-\alpha)) \quad (61)$$

e

$$t \pm v_L^{1/2} z^*((R+1)\alpha) \quad (62)$$

onde

$$z^* = (t^* - t) / v_L^{1/2} \quad (63)$$

e R é o número de replicações bootstrap.

3.5 Análise de Medidas Repetidas em Planejamentos Crossover 2x2

Este capítulo tem como objetivo principal apresentar uma revisão sobre a análise de medidas repetidas no contexto de ensaios de bioequivalência pelo fato de se tratar de um assunto ainda não muito explorado e a disponibilidade de material escasso. Serão apresentados alguns tópicos importantes, como a escolha de uma adequada estrutura de covariância, entre outros, para assim explanar um modelo simples, visto que a análise de medidas repetidas é um assunto complexo e vasto e seria necessário um trabalho mais extenso para especificar todos detalhes, aplicações e derivações.

3.5.1 Caracterização

Medidas repetidas são investigações científicas em que o interesse é dirigido para a avaliação do comportamento de uma ou mais variáveis respostas ao longo de uma dimensão específica como o tempo, à distância de uma certa origem ou dosagem de uma determinada substância. De uma forma geral, esse tipo de pesquisa envolve observações de um conjunto de unidades de investigação classificados em diferentes subpopulações, segundo um ou mais fatores ou tratamentos ao longo de diversas condições de avaliação.

A análise clássica de medidas repetidas utiliza modelos lineares gerais, nas duas soluções tradicionais, caracterizadas como análise univariada de perfis e análise multivariada de perfis. Também se inclui a análise de tendências de variação, onde pode-se investigar a existência de tendências lineares, quadráticas ou de graus superiores, e dentro dessas tendências, analisar as possíveis diferenças entre os tratamentos. A análise moderna de medidas repetidas utiliza modelos mistos para acomodar a estrutura de correlação produzida devido às medidas repetidas.

De acordo com Jones & Kenward (1990) pode-se dizer que o delineamento crossover é um exemplo típico de seqüência de observações ou medidas repetidas, já que consiste em observações de sucessivos períodos de tratamento. Entretanto, uma seqüência de observações pode também ser coletada enquanto o sujeito está sob o mesmo tratamento experimental. Usa-se o termo medida repetida para referir-se a essa seqüência.

Na maioria dos estudos é feita uma tentativa para assegurar que as medidas sejam feitas em tempos equivalentes para cada sujeito. Se todos os sujeitos entram no experimento ao mesmo tempo, então as medidas também poderão ser feitas aproximadamente no mesmo tempo absoluto.

Da mesma maneira, os tempos serão os mesmos relativos ao mesmo ponto inicial. As medidas repetidas podem conter informações úteis na forma de verificar a existência de efeitos residuais. É possível que um efeito possa aumentar durante o período de tratamento e isto poderá ser mostrado completamente na análise de medidas repetidas na forma da interação efeito residual por tempo.

As medidas repetidas de cada sujeito podem ser reduzidas a poucas estatísticas descritivas, que podem ser analisadas separadamente usando técnicas univariadas padrão ou técnicas multivariadas análogas. O uso de estatísticas descritivas tem as grandes vantagens de simplicidade e robustez, além da maior facilidade de compreensão. Isso é aplicável sempre que é possível resumir aspectos relevantes de perfis com poucas estatísticas escolhidas a priori.

A dificuldade na análise de medidas repetidas é a escolha adequada da estrutura de covariância, devido à natureza seqüencial dos dados em cada unidade experimental. Se é possível assumir que a estrutura de covariância é uniforme, então os métodos clássicos de análise, como a análise de variância univariada poderão ser usados, caso contrário métodos mais sofisticados devem ser empregados.

3.5.2 Estrutura de Covariância

Na prática existem três tipos de análise estatística comumente usadas na análise de medidas repetidas. A análise de variância univariada é como se fosse um experimento em parcela subdividida (split-plot), com os tratamentos (drogas) como o fator da parcela e o tempo como fator da subparcela. Esse tipo de análise ignora a estrutura de covariância, o que pode resultar em conclusões incorretas. As análises univariada e multivariada para transformações lineares das medidas repetidas, tais como as inclinações e outras tendências em curvas de regressão, diferenças entre respostas de diferentes pontos de tempo, evitam a estrutura de covariância, podendo resultar em análises ineficientes, o que seria equivalente a um desperdício de dados. E os métodos baseados em modelos mistos com estruturas de covariância nos permitem eficientemente conduzir a questão diretamente na modelagem da estrutura de covariância.

A justificativa para a maioria das análises serem baseadas na Anova padrão podem ser encontradas no processo de aleatorização usado no delineamento experimental. Independentemente de qual delineamento seja usado, e se os efeitos residuais serão ou não ajustados, a alocação aleatória dos sujeitos nas seqüências pode ser usada para proporcionar

algum suporte de igualdade nas estruturas de covariância entre seqüências, mas não como se têm visto, para propriedades de estruturas dentro das seqüências. Uma forte restrição na aleatorização do planejamento ou a inclusão de efeitos residuais na análise significa que não se pode justificar uma particular estrutura de covariância dentro de seqüências de um planejamento isoladamente.

Normalmente começa-se testando uma particular estrutura de covariância para modelar a estrutura dos dados, e após essa tentativa inicial, testam-se outras estruturas de covariância, visando construir um modelo mais adequado aos dados.

3.5.3 Metodologia de Modelos Mistos para Análise de Medidas Repetidas de Dados em Estudos Crossover

A metodologia de modelos mistos constitui-se atualmente no procedimento mais utilizado e difundido para a avaliação de medidas repetidas, resultando em uma generalização do modelo linear geral, separando os efeitos fixos dos aleatórios, cuja caracterização é dada por:

$$Y = X\beta + Zu + \varepsilon, \quad (64)$$

onde Y é o vetor de dados observados, X é a matriz de planejamento para os efeitos fixos, β é o vetor de parâmetros associados aos efeitos fixos, u é o vetor de parâmetros associados aos efeitos aleatórios e ε é o vetor de erros aleatórios não-observáveis. Os vetores u e ε são não correlacionados, com esperanças nulas e matrizes de covariância G e R , respectivamente.

A parametrização das estruturas de covariância vai desde a estrutura mais restritiva, que é a de Componentes Simétricos, até a Não-Estruturada, que é a estrutura menos exigente ou restritiva, passando por várias outras, como por exemplo, Auto-Regressivo de 1ª Ordem, Toeplitz, H-F, Componentes de Variância, Toeplitz Heterogêneo de 1ª Ordem, Auto-Regressivo de 1ª Ordem e Média Móvel e Antedependência de 1ª Ordem.

O modelo estatístico apresentado na seção 3.2.2 será usado para representar a análise de medidas repetidas de um delineamento crossover. Serão acrescentadas a esse modelo as interações simples entre as formulações e o tempo e o período e o tempo. Essa nova proposta de ajuste de modelo é dada por

$$Y_{ijkl} = \mu + S_{ik} + P_j + F_{(j,k)} + T_l + T_l * F_{(j,k)} + T_l * P_j + R_{(j-1,k)} + \varepsilon_{ijkl} \quad (65)$$

onde

Y_{ijkl} : variável resposta (por exemplo, $ASCt_k$, Cmax) do i-ésimo indivíduo na k-ésima seqüência no j-ésimo período e l-ésimo tempo;

μ : média geral;

S_{ik} : efeito do i-ésimo sujeito na k-ésima seqüência, onde i (sujeito) = 1, 2, 3, ..., n_k e k (seqüência) = 1, 2, ..., g;

P_j : efeito fixo do j-ésimo período, $j = 1, 2, \dots, p$ tal que $\sum_j P_j = 0$;

$F_{(j,k)}$: efeito fixo da formulação administrada na k-ésima seqüência e j-ésimo período tal que, $\sum F_{(j,k)} = 0$;

T_l : efeito do l-ésimo tempo como medida repetida, onde $l = 1, 2, \dots, t$;

$T_l * F_{(j,k)}$: interação simples entre o l-ésimo tempo e as formulações administradas na k-ésima seqüência e j-ésimo período;

$T_l * P_j$: interação simples entre o l-ésimo tempo e o j-ésimo período;

$R_{(j-1,k)}$: efeito residual fixo de primeira ordem da formulação administrada na k-ésima seqüência e (j-1)-ésimo período, tal que $R_{(0,k)} = 0$ e $\sum R_{(j-1,k)} = 0$;

ε_{ijkl} : erro aleatório (intra-sujeito).

Assume-se que $\{S_{ik}\}$ e $\{\varepsilon_{ijkl}\}$ são independentes, identicamente distribuídos (i.i.d) com média zero e variâncias σ_s^2 e σ_t^2 respectivamente, e mutuamente independentes.

A esse modelo pode-se também incluir os efeitos residuais de segunda ordem e a interação simples entre efeito residual e os períodos de tempo, para verificar a ocorrência de um aumento nos efeitos residuais durante o período de tratamento.

O ajuste do modelo foi realizado no programa SAS. O programa 1 mostra o ajuste para a variável concentração, na opção "Repeated Measures", onde a estrutura de covariância escolhida foi a CS (Componentes Simétricos) e o programa 2, realizado em outra opção do SAS, "Linear

Models”, traz o ajuste para todas as medidas farmacocinéticas envolvidas no estudo. Abaixo segue a rotina de programação.

```
PROGRAMA 1

*** Repeated Measures ANOVA Analysis ***;
proc mixed data=ParteMED.medida_repetida;
  class SUJEITO TEMPO DROGA SEQUENCIA PERIODO;
  model CONCENT = TEMPO DROGA SEQUENCIA PERIODO
DROGA*TEMPO TEMPO*PERIODO/ ddfm=satterth
  htype=3 outp=WORK._PRED;
  repeated TEMPO / sub=SUJEITO(SEQUENCIA) type=cs;
run;

PROGRAMA 2

*** Linear Models Analysis ***;
proc glm data=PARTEMED.resumo;
  class SUJEITO PERIODO SEQUENCIA DROGA;
  model TMAX CMAX KE T1_2 ASCT ASCINF = PERIODO SEQUENCIA
DROGA
  SUJEITO(SEQUENCIA) / SS3;
  test H=sequencia E=sujeito(sequencia);
```

3.6 Algumas Considerações sobre Bioequivalência

Este capítulo foi baseado no Manual da ANVISA, e tem como maior intuito fornecer uma visão geral de alguns aspectos muito importantes em ensaios de bioequivalência que não foram abordados nos capítulos anteriores.

3.6.1 Observações Atípicas (“Outliers”) em Estudos de Bioequivalência

Um dos problemas geralmente encontrados nos estudos de biodisponibilidade e bioequivalência é o fato do conjunto de dados conter valores extremamente altos ou

extremamente baixos. Essas observações atípicas (outliers) podem representar um grande impacto na avaliação da bioequivalência. Existem basicamente três possíveis tipos de outliers nos estudos de bioequivalência:

1. Valores inesperados na curva de concentração plasmática em função do tempo, para determinados tempos de coleta;
2. Valores extremamente altos ou baixos para determinada formulação (teste ou referência) e;
3. Indivíduos incomuns que exibem biodisponibilidade muito inferior ou muito superior com relação ao medicamento referência, ou seja, uma diferença muito grande para o comportamento da curva entre as duas formulações, implicando em valores diferenciados para todas as medidas farmacocinéticas avaliadas.

Para o primeiro tipo de outlier, as observações inesperadas da curva de concentração plasmática em função do tempo, normalmente têm um efeito pequeno sobre o cálculo de ASC (área sobre a curva de concentração), e conseqüentemente tem um pequeno efeito na comparação da biodisponibilidade.

O segundo tipo de outlier caracteriza-se pela presença de valores muito altos ou muito baixos para a medida avaliada em determinada formulação, ou em outras palavras, significa que a distância entre os valores da medida farmacocinética para o medicamento teste e o medicamento referência de um determinado indivíduo é muito grande. Esse tipo de outlier tem certamente um grande impacto na avaliação comparativa, considerando o critério da bioequivalência média, pelo fato da média ser um estimador muito sensível à presença de valores extremos.

Para o terceiro tipo de outlier, similarmente ao segundo caso, a presença de um indivíduo atípico pode vir a anular a conclusão de bioequivalência do estudo. A constatação de biodisponibilidades extremamente baixas, ou extremamente altas entre os voluntários, pode vir a indicar que a variabilidade da resposta a alguma das formulações não é homogênea.

Portanto, uma análise comparativa da biodisponibilidade, sob a presença de indivíduos potencialmente atípicos, possivelmente pode ter que lidar com a rejeição da bioequivalência entre as formulações testadas, quando na realidade, elas são bioequivalentes. Nesse sentido, a seleção e

os critérios para inclusão de voluntários nos estudos podem atuar como fatores determinantes no intuito de evitar a presença de voluntários atípicos nos ensaios de bioequivalência.

Para um determinado conjunto de dados, um indivíduo outlier potencial pode ser detectado através da plotagem dos valores da medida farmacocinética avaliada para os diversos voluntários, da formulação teste contra referência. O indivíduo é detectado em função do seu respectivo ponto apresentar um grande desvio em relação à reta $y = x$.

Uma outra ferramenta visual bastante útil para a identificação de possíveis outliers é o diagrama de caixas ou boxplot, construído de forma a fornecer informações sobre a simetria e variabilidade dos dados.

Os estudos para detecção de valores potencialmente atípicos têm recebido bastante atenção nas últimas décadas, e em particular, o desenvolvimento de métodos para o tratamento de outliers em modelos de regressão linear. Como normalmente os estudos de biodisponibilidade não se aplicam à utilização de modelos de regressão linear, torna-se necessário o desenvolvimento de novos métodos que venham a proporcionar maneiras de identificar observações atípicas nesse caso. Uma vez detectada uma observação atípica, a questão de interesse, portanto, é saber qual o peso daquele valor em relação à avaliação da bioequivalência. Com base nas indicações dos órgãos reguladores e nas legislações vigentes, a retirada de voluntários de um estudo em função de seus dados não se apresentarem de maneira similar aos dos outros voluntários sem justificativa convincente pode afetar a validação do estudo. Em muitos casos, não há como precisar se essa aparente não similaridade é resultado de erros laboratoriais, erros de transcrição ou outras causas não relacionadas à bioequivalência. Portanto, a retirada de voluntários do estudo não é recomendada, especialmente para experimentos não replicados. A legislação brasileira preconiza que no caso de retirada de voluntários que apresentarem comportamento discrepante nas medidas de absorção, em relação aos demais voluntários, sua exclusão do estudo deverá ser justificada. Nesse caso, deverão ser apresentados os resultados do estudo com e sem a inclusão de seus dados, no sentido de avaliar o impacto da retirada dessa observação.

3.6.2 Uso de Covariáveis

Normalmente em ensaios clínicos há a necessidade de usar informações adicionais analisadas sobre cada sujeito. Esses dados adicionais são chamados covariáveis e podem ser colocados sobre dois importantes aspectos (usos). Pode ser que deseja-se saber se o efeito de tratamento é relacionado com os valores avaliados para uma covariável, ou ainda, se a covariável não parece ser relacionada com o efeito de tratamento.

A covariável pode ser categórica ou contínua. A introdução de uma covariável categórica é equivalente a introduzir um fator a mais no ensaio, e assim simplesmente necessita-se generalizar a análise de variância. Já a introdução de uma covariável contínua no modelo, deixa a análise bem mais complexa, e decomposta em duas tabelas: análise de covariância dos sujeitos totais e análise de covariância das diferenças dentro-sujeitos. Mais detalhes em Jones & Kenward (1990).

3.6.3 Dados Faltantes (“Missings Values”)

Os valores “missings” acontecem quando há pares de observações faltantes. Suponha que no grupo i há n_i sujeitos com pares completos de observações (Y_{1ik}, Y_{2ik}) , e m_{ij} sujeitos com uma única observação Z_{ijk} no j -ésimo período ($j = 1, 2$; $i = 1, 2$).

Há três formas para se analisar esses dados. A primeira e mais simples é analisar apenas os pares completos. Desde que os dados remanescentes contêm apenas a informação entre-sujeitos ou entre-indivíduos, o efeito de descartar essas observações é provavelmente maior para o teste da interação tratamento (droga) por período. Este teste é geralmente de menor poder, então qualquer possível perda de sensibilidade é indesejável.

O segundo método de análise foi sugerido por (Patel (1985) Apud Jones & Kenward (1990)) e faz uso da simplificação que ocorre quando as observações são faltantes apenas do segundo período, isto é, $m_{12} = m_{22} = 0$. Usa-se uma matriz de dispersão incondicional para os pares (Y_{1ik}, Y_{2ik}) ,

$$V \left\{ \begin{bmatrix} Y_{1ik} \\ Y_{2ik} \end{bmatrix} \right\} = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \quad (66)$$

e considerando $E[Y_{ijk}] = E[Z_{ijk}] = \mu_{ij}$, é possível, com uma reparametrização, particionar a verossimilhança dos dados em duas componentes:

$$l(y_1, z_1; \mu_{11}, \mu_{21}, \sigma_{11}) \times l(y_2 / y_1; \mu_{12}^*, \mu_{22}^*, \beta, \sigma_{2.1}) \quad (67)$$

onde

$$y_1 = (y_{111}, y_{112}, \dots, y_{21n_2})^t \quad (68)$$

$$z_1 = (z_{111}, z_{112}, \dots, z_{21m_2})^t \quad (69)$$

$$\mu_{i2}^* = \mu_{i2} - \beta \mu_{i1} \quad (70)$$

$$\beta = \sigma_{12} / \sigma_{11} \quad (71)$$

e

$$\sigma_{2.1} = \sigma_{22} - \beta \sigma_{12} \quad (72)$$

Desde que nenhuma das novas reparametrizações ocorra em ambas componentes da verossimilhança, as duas componentes podem ser maximizadas separadamente. De procedimentos padrões, obtêm-se da primeira componente:

$$\hat{\mu}_{i1} = (n_i + m_{i1})^{-1} (y_{i1} + z_{i1}) = p_i \bar{y}_{i1} + (1-p) \bar{z}_{i1}, \quad i = 1, 2 \quad (73)$$

onde

$$p_i = \frac{n_i}{n_i + m_{i1}} \quad (74)$$

e

$$\hat{\sigma}_{11} = s_{11} = \frac{1}{n_1 + n_2 + m_{11} + m_{21} - 2} * \sum_{i=1}^2 \left[\sum_{k=1}^{n_i} (y_{ik} - \hat{\mu}_{i1})^2 + \sum_{k=1}^{m_{i1}} (z_{ik} - \hat{\mu}_{i1})^2 \right]. \quad (75)$$

Da segunda componente,

$$\hat{\beta} = s_{12}^{(y)} / s_{11}^{(y)} \quad (76)$$

Esses podem ser estimados ao substituir as estimativas dadas pelos parâmetros desconhecidos.

A análise é então baseada nos contrastes de $\hat{\mu}_{ij}$. Tem-se a interação direta por período:

$$\hat{\lambda} = \frac{1}{2} [\hat{\mu}_{11} + \hat{\mu}_{12} - \hat{\mu}_{21} - \hat{\mu}_{22}] \quad (86)$$

e o efeito direto de tratamento por

$$\hat{\tau} = \frac{1}{4} [\hat{\mu}_{11} - \hat{\mu}_{12} - \hat{\mu}_{21} + \hat{\mu}_{22}]. \quad (87)$$

Os erros padrões destes estimadores são obtidos das variâncias e covariâncias estimadas de $\hat{\mu}_{ij}$ e estatísticas convencionais, ou seja, $\hat{\lambda}/[E.P(\hat{\lambda})]$. Patel (1985) Apud Jones & Kenward (1990), sugere usar a distribuição t de Student com graus de liberdade baseados no número de pares completos ($n_1 + n_2 - 2$), como uma aproximação para o teste destas estatísticas. O autor mostra em um estudo de simulação que essa aproximação é adequada em ensaios de pequena dimensão, exceto quando a proporção de pares incompletos é grande (maior que 40%), uma vez que nesses casos os resultados apresentam validade duvidosa.

O terceiro método de análise aplica-se quando as observações são faltantes em ambos os períodos. As estimativas são obtidas separadamente, ou seja, dos dados completos e dos dados com os “missings”, e então combinadas usando os inversos das estimativas das variâncias como pesos.

Assumindo que a variância é a mesma em ambos períodos, pode-se escrever:

$$\Sigma = \begin{bmatrix} \sigma^2 + \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma^2 + \sigma_s^2 \end{bmatrix} \quad (88)$$

e, então,

$$V[\hat{\lambda}_C] = \frac{1}{2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) (2\sigma_s^2 + \sigma^2) = K_C (2\sigma_s^2 + \sigma^2) \quad (89)$$

e

$$V[\hat{\lambda}_I] = \frac{1}{4} \left(\frac{1}{m_{11}} + \frac{1}{m_{12}} + \frac{1}{m_{21}} + \frac{1}{m_{22}} \right) (\sigma_s^2 + \sigma^2) = K_I (\sigma_s^2 + \sigma^2). \quad (90)$$

Combinando as duas estimativas, obtêm-se

$$\hat{\lambda} = \frac{K_C^{-1} (2\sigma_s^2 + \sigma^2)^{-1} \hat{\lambda}_C + K_I^{-1} (\sigma_s^2 + \sigma^2)^{-1} \hat{\lambda}_I}{K_C^{-1} (2\sigma_s^2 + \sigma^2)^{-1} + K_I^{-1} (\sigma_s^2 + \sigma^2)^{-1}} = \frac{K_I \hat{\lambda}_C + (1 + \rho^2) K_C \hat{\lambda}_I}{K_I + (1 + \rho^2) K_C}, \quad (91)$$

onde

$$\rho^2 = \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2} \quad (92)$$

é a correlação intra-classe.

A variância de $\hat{\lambda}$ é aproximadamente

$$V[\hat{\lambda}] = \sigma^2 \frac{K_C K_I \rho^2 (1 + \rho^2)}{K_C (1 + \rho^2) + K_I \rho^2}. \quad (93)$$

Uma aproximação para os graus de liberdade do estimador combinado pode ser obtida usando a expressão de Satterhwaite. (Satterhwaite (1946) Apud Jones & Kenward (1990)).

3.6.4 Bioequivalência Individual e Populacional

O trabalho abordou toda a análise estatística baseada no critério da bioequivalência média, mas é importante ressaltar a existência dos conceitos relativos a bioequivalências individual e populacional. Esses novos critérios para a avaliação da bioequivalência podem ser bastante úteis em alguns casos. De acordo com a definição apresentada nos capítulos anteriores, a bioequivalência média focaliza somente a comparação das médias populacionais referentes às medidas farmacocinéticas de interesse. Em contrapartida, os critérios da bioequivalências

individual e populacional englobam além das médias, as respectivas variâncias associadas às medidas farmacocinéticas avaliadas.

O critério da bioequivalência populacional leva em consideração a variabilidade total associada à medida de interesse. Já o critério da bioequivalência individual agrega a variabilidade intra-individual nas formulações teste e referência, além do componente de variação associado à interação entre indivíduos e formulação, que equivale à variação entre as médias das formulações T e R devido às diferenças existentes entre os indivíduos.

4. EXEMPLO

A ilustração dos métodos para análise de bioequivalência média entre duas formulações R (referência) e T (teste) será realizada neste trabalho utilizando dados hipotéticos retirados do Manual de Boas Práticas em Biodisponibilidade e Bioequivalência da ANVISA, devido ao sigilo e importância que tais dados carregam consigo. As concentrações de 24 voluntários são apresentadas nas Tabelas 6 e 7, com coletas realizadas em 13 tempos diferentes (dados longitudinais), respectivamente nas formulações R e T.

As Tabelas 8 e 9 mostram as principais medidas farmacocinéticas calculadas para as concentrações das Tabelas 6 e 7 de acordo com o planejamento de estudo crossover, considerando a seqüência de administração dos medicamentos (RT=1, TR=2) e também o período de administração (períodos 1 e 2) das formulações R e T, respectivamente. As medidas farmacocinéticas selecionadas nas Tabelas 8 e 9 são o tempo onde a concentração máxima foi alcançada, a concentração máxima observada, a constante de velocidade de eliminação total do fármaco no organismo, a meia-vida de eliminação do fármaco, a área sobre a curva de concentração do tempo zero ao tempo t e a área sobre a curva de concentração do tempo zero ao tempo infinito, de acordo com a ordem das colunas.

Tabela 6. Concentrações plasmáticas (ng/ml) ao longo do tempo (horas) referentes ao medicamento de referência (R).

Sujeito	Tempo (horas)												
	0	0,5	1	1,5	2	2,5	3	3,5	4	6	8	10	12
1	0	0	18,3	82,9	205,4	62,3	53,9	33,6	19,8	0	0	0	0
2	0	19,2	23,4	104,7	287	275,9	217,9	133,2	77,5	36	17,9	14,5	0
3	0	0	9,4	0	23,5	132,5	198,3	135,9	81,3	25,4	0	0	0
4	0	0	0	115,2	193,7	204,2	150,5	142,2	58,4	36,2	17,6	0	0
5	0	0	0	25,8	56	270,3	115,4	112,3	38,6	13	0	0	0
6	0	0	70,3	129,6	60,7	51,9	32,6	17,8	0	0	0	0	0
7	0	29,2	256,9	92,9	55,2	32	28,2	22,9	0	0	0	0	0
8	0	19,1	233,3	60,26	103,3	88,4	62,9	34,5	33	28,2	0	0	0
9	0	0	32,3	215,9	317,2	153,9	99,5	92,3	36,9	12,6	0	0	0
10	0	13,2	12	98,5	259,2	113,6	51,9	45	30,9	21,8	7,2	0	0
11	0	0	0	81,1	117,6	64,3	63,9	58,2	41,7	15,2	19	10,9	0
12	0	0	0	114,2	138,4	146,6	156,2	177,9	80,5	37,2	22,8	17,5	0
13	0	0	94,5	125,8	154,2	137,9	95,8	39,8	28,7	0	0	0	0
14	0	32,5	114,2	171,6	123,9	91,9	84	74,3	50,2	19,7	16,5	15,9	16,5
15	0	0	0	22,8	82,5	178,3	58,2	44,3	24,8	11,9	8,4	0	0
16	0	19,5	54,2	97,2	165,3	157,9	145,6	60,6	29,2	22,6	15,9	0	0
17	0	23,5	59,6	145,2	131,5	80,9	64,1	61,8	28,9	14,4	18,5	0	0
18	0	0	19,2	39,5	115,2	297,4	164,2	145,2	80,5	32,2	11,5	0	0
19	0	0	78,2	209,7	172,6	159,1	66,9	37,2	29,4	9,4	0	0	0
20	0	11,5	63,9	248,6	186,6	74,2	29,5	23,4	15	0	0	0	0
21	0	0	72,9	215,3	168,2	51,9	37,8	33,9	15,7	0	0	0	0
22	0	0	357,1	372,3	177,4	122,3	64,9	55,6	14,9	0	0	0	0
23	0	0	209,6	515,5	230,8	102	129,4	64,9	21,4	25,2	20,9	10	0
24	0	0	25,3	183,1	70,3	25,1	14,2	11,5	8,4	0	0	0	0

Tabela 7. Concentrações plasmáticas (ng/ml) ao longo do tempo (horas) referentes ao medicamento de teste (T).

Sujeito	Tempo (horas)												
	0	0,5	1	1,5	2	2,5	3	3,5	4	6	8	10	12
1	0	0	29,3	75,2	215,9	48,2	26,9	19,5	14,3	0	0	0	0
2	0	0	67,3	315,8	335,5	209,8	233,8	116,9	58,7	28,4	22,1	15,2	0
3	0	0	89,5	105,3	285,5	180,9	65,4	32,9	23	15,9	12,5	0	0
4	0	0	0	196,9	291,4	297,7	324,5	97,2	56,9	23,5	18,7	0	0
5	0	0	13,7	188,8	221,6	147,2	58,4	44,3	29,9	11,9	0	0	0
6	0	0	75,8	188	82,3	63,7	48	11,3	10,7	0	0	0	0
7	0	129,6	291,4	209,2	85,9	54,7	32,5	30,9	28,4	12,4	9	0	0
8	0	39,5	250,4	226,9	80,4	75,3	59,2	49,1	42,5	0	0	0	0
9	0	0	12,5	256,3	304,8	219,7	115	33,1	14,9	15,3	0	0	0
10	0	12,2	29,4	12,4	127,8	65,8	81,7	49,3	39,6	33,1	15,2	0	0
11	0	49,5	117	103,5	78,3	74,3	63,2	55,1	48,7	39,2	23	12,2	0
12	0	0	0	125,5	175	208,5	161,3	95,9	65	18	13,7	0	0
13	0	0	75,2	141,9	95,4	72,9	56	41	24,7	0	0	0	0
14	0	37,7	107,5	215,2	100,3	96,5	64,6	56,5	51,7	31,9	23,2	17,1	0
15	0	0	43,4	93,3	117,7	91	42,6	36,9	27,2	12,9	0	0	0
16	0	25,5	75,5	203,9	177,3	88,5	53,5	29,9	27	25,8	16	0	0
17	0	43,4	45,2	78,2	205,3	95,8	45,7	32,5	19,3	17,8	16,5	12,2	0
18	0	0	23,4	72,4	256,3	166,9	124,8	96,2	88,2	25,4	12,9	0	0
19	0	0	0	227	195,2	82,5	45,2	22,9	18	11,2	0	0	0
20	0	9,3	32,5	189,4	277,5	83,9	77,4	53,6	33,1	12,4	8,9	0	0
21	0	0	45,3	133,2	267,4	165,2	95,4	52,4	33,8	15,4	0	0	0
22	0	0	139,3	325,9	249,3	127,2	104,8	64,2	11,9	0	0	0	0
23	0	22,3	245,8	586,2	241,2	105	100,2	42,5	20,4	15,8	9,6	0	0
24	0	44,8	250,1	53,6	38,9	22,5	23,7	18,9	14,2	0	0	0	0

Tabela 8. Medidas Farmacocinéticas dos sujeitos referentes ao medicamento de referência (R).

Sujeito	período	seqüência	Tmax	Cmax	Ke	t1/2	ASCt	ASCinf
1	1	RT	2	205,4	1	0,69	252,95	252,92
2	2	TR	2	287	0,41	1,69	764,33	785,23
3	1	RT	3	198,3	0,68	1,03	402,23	414,41
4	2	TR	2,5	204,2	0,3	2,31	583,5	624,6
5	1	RT	2,5	270,3	0,84	0,83	364,6	367,1
6	2	TR	1,5	129,6	1,07	0,65	181,45	193,63
7	2	TR	1	256,9	0,33	2,07	258,65	321,37
8	1	RT	1	233,3	0,08	8,66	448,55	772,56
9	2	TR	2	317,2	0,78	0,88	526,88	530,35
10	1	RT	2	259,2	0,37	1,86	393,33	405,44
11	2	TR	2	117,6	0,27	2,53	334,88	363,72
12	1	RT	3,5	177,9	0,19	3,68	622,28	697,6
13	1	RT	2	154,2	1,12	0,62	359,88	356,86
14	2	TR	1,5	171,6	0,24	2,95	529,65	599,8
15	1	RT	2,5	178,3	0,27	2,56	264,65	287,29
16	1	RT	2	165,3	0,15	4,56	473,65	562,38
17	2	TR	1,5	145,2	0,36	1,92	385,23	417,87
18	2	TR	2,5	297,4	0,49	1,42	578,38	590,51
19	2	TR	1,5	209,7	0,56	1,25	417,4	424,91
20	1	RT	1,5	248,6	1,21	0,57	337,6	334,99
21	2	TR	1,5	215,3	1,04	0,67	312,78	312,16
22	2	TR	1,5	372,3	1,15	0,6	593,43	591,43
23	1	RT	1,5	515,5	0,23	3	765,05	798,33
24	1	RT	1,5	183,1	0,53	1,32	175,25	182,85

Tabela 9. Medidas Farmacocinéticas dos sujeitos referentes ao medicamento de teste (T).

Sujeito	período	seqüência	Tmax	Cmax	Ke	t1/2	ASCt	ASCinf
1	2	RT	2	215,9	0,63	1,1	225,38	233,71
2	1	TR	2	335,5	0,16	4,44	844,33	926,39
3	2	RT	2	285,5	0,15	4,55	465,3	534,8
4	1	TR	3	324,5	0,35	1,96	759,38	793,66
5	2	RT	2	221,6	0,53	1,32	398,18	408,92
6	1	TR	1,5	188	1,18	0,59	247,93	246,33
7	1	TR	1	291,4	0,28	2,46	495,4	518,33
8	2	RT	1	250,4	0,33	2,09	443,53	529,26
9	1	TR	2	304,8	0,82	0,85	519,93	523,34
10	2	RT	2	127,8	0,27	2,57	335,4	376,48
11	1	TR	1	117	0,29	2,38	480,13	509,73
12	2	RT	2,5	208,5	0,52	1,33	527,75	540,35
13	2	RT	1,5	141,9	0,66	1,05	272,08	284,84
14	1	TR	1,5	215,2	0,16	4,45	548,18	640,77
15	2	RT	2	117,7	0,4	1,72	272,25	291,31
16	2	RT	1,5	203,9	0,13	5,52	444,4	555,81
17	1	TR	2	205,3	0,07	9,55	390,18	546,06
18	1	TR	2	256,3	0,48	1,46	556,85	571,03
19	1	TR	1,5	227	0,27	2,55	331,3	361,26
20	2	RT	2	277,5	0,44	1,57	445,78	457,09
21	1	TR	2	267,4	0,46	1,5	452,7	470,42
22	1	TR	1,5	325,9	1,19	0,58	520,23	518,33
23	2	RT	1,5	586,2	0,19	3,68	747,9	789,24
24	2	RT	1	250,1	0,51	1,35	244	257,24

4.1 Análise Preliminar do Exemplo usando variável $ASC t_k$

Para uma análise preliminar dos dados do exemplo serão utilizadas as Tabelas 8 e 9. Inicialmente faz-se uma análise descritiva da área sobre a curva de concentração do tempo zero ao tempo t $ASC t$, sem separar os tratamentos referência e teste. As médias dessas áreas estão sintetizadas na Tabela 10.

Tabela 10. Médias das áreas sobre as curvas de concentração do tempo 0 ao tempo t .

	Período 1	Período 2	
Seqüência RT	405,0017	401,8292	403,4155
Seqüência TR	512,2117	455,5467	483,8792
	458,6067	428,6880	443,6474

Pode-se ver pela Tabela 10 que os resultados indicam que as médias amostrais das $ASC t_k$ para as formulações de teste e referência são maiores para a seqüência 2 (TR). Para um delineamento crossover 2x2, a interação entre período e seqüência representa o efeito de droga se não existem efeitos residuais. Portanto, um teste preliminar para a verificação da existência de efeitos residuais deve ser necessariamente realizado antes da avaliação da bioequivalência entre as formulações. Também é recomendável checar o ajuste dos dados à normalidade. A variável $ASC t$ possui um bom ajuste à distribuição normal (Anderson Darling: p-valor = 0,204 e Kolmogorov-Smimov: p-valor > 0,15).

4.1.1 Análise dos Efeitos Residuais

Para o cálculo dos efeitos residuais de cada seqüência monta-se uma tabela com a soma da medida farmacocinética em estudo nos períodos 1 e 2. Após as somas, é calculada a média para as k seqüências ($k = 1, 2$). Com base nessas somas faz-se o cálculo da variância amostral dos efeitos residuais, mediante o quadrado das diferenças entre a soma das medidas e a média para cada seqüência, indicada no item 3.2.3.1. Os cálculos resumidos encontram-se na Tabela 11.

Tabela 11. Somas das $ASCt$ para as k seqüências.

Seqüência 1 (RT)	Seqüência 2 (TR)	desv SEQ1	desv SEQ2
478,33	1608,66	107912,78	410754,95
867,53	1342,88	3684,3929	140716,27
762,78	429,38	1940,473	289851,23
892,08	754,05	7267,4261	45671,252
728,73	1046,81	6099,735	6249,1661
1150,03	815,01	117785,69	23332,053
631,96	1077,83	30579,797	12115,772
536,9	775,41	72862,637	36997,881
918,05	1135,23	12369,71	28046,759
783,38	748,7	549,94002	47986,553
1512,95	765,48	498604,32	40916,524
419,25	1113,66	150218,88	21287,296
Soma = 9681,97	Soma = 11613,1	1009875,8	1103925,7
Média = 806,8308	Média = 967,7583		

Da tabela, vemos que o efeito residual é estimado com base na diferença entre as duas seqüências ($967,7583 - 806,8308 = 160,9275$), com variância estimada calculada a partir da fórmula (11) da seção 3.2. O cálculo é dado por

$$Var(\hat{R}) = \hat{\sigma}_u^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) = 96.081,885 * (1/12 + 1/12) = 16.013,648.$$

E, assim, a estatística de teste resulta em

$$T_r = \frac{\hat{R}}{\sqrt{(\hat{\sigma}_u^2) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = 160,9275/126,54504 = 1,2717,$$

que ao ser comparada com o valor da estatística padrão t de Student, com 22 graus de liberdade e nível de significância 0,05 ($t_{\alpha/2}(n_1 + n_2 - 2) = 2,074$) nos leva a aceitar a hipótese de nulidade. Em outras palavras, pode-se dizer que não há evidências estatísticas suficientes para concluir quanto à existência de efeitos residuais a um nível de significância de 5% (p-valor = 0,22). A inexistência de efeito residual não implica necessariamente que tais efeitos sejam nulos, mas que se existirem, têm a mesma intensidade em ambas as seqüências de tratamentos. Quando há

inexistência de efeitos residuais, os efeitos de droga (tratamentos) podem ser estimados com base nos dados de ambos os períodos.

4.1.2 Análise do Efeito de Tratamento

Para o cálculo do efeito de droga monta-se uma tabela com as diferenças das medidas farmacocinéticas divididas por 2 nos períodos 1 e 2. Após as diferenças são calculadas as médias para as k seqüências ($k = 1, 2$). Com base nessas diferenças faz-se o cálculo da variância amostral dos efeitos das drogas, usando o quadrado das diferenças, como feito no item 4.2. Os cálculos resumidos encontram-se na Tabela 12.

Tabela 12. Diferenças das ASC_t para as k seqüências.

Seqüência 1 (RT)	Seqüência 2 (TR)	desv_SEQ1	desv_SEQ2
13,785	-40	148,8095	136,1306
-31,535	-87,94	1097,0172	3553,054
-16,79	-33,24	337,68656	24,08356
2,51	-118,375	0,8533141	8107,652
28,965	3,475	749,59595	1011,717
47,265	-72,625	2086,5482	1961,826
43,9	-9,265	1790,4534	363,5696
-3,8	-2,475	29,011689	668,6103
14,625	10,765	170,009	1528,615
-54,09	43,05	3099,8448	5095,461
8,575	-69,96	48,842627	1732,849
-34,375	36,6	1293,2115	4216,23
Soma = 19,035	Soma = -339,99	10851,884	28399,8
Média = 1,58625	Média = -28,3325		

Da tabela, vemos que a estimativa para o efeito da droga é a diferença entre as médias das seqüências 1 e 2, isto é,

$$\hat{F} = \bar{d}_1 - \bar{d}_2 = \bar{Y}_T - \bar{Y}_R = 1,58625 - (-28,3325) = 29,91875,$$

e a variância amostral é calculada com base em (19), resultando em 1.784,1673. A estatística de teste é dada por:

$$T_d = \frac{\hat{F}}{\sqrt{(\hat{\sigma}_d^2) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = 29,91875/17,24416467 = 1,7350$$

que nos leva a aceitar a hipótese nula ao comparar com o valor tabelado t de Student com 22 graus de liberdade e nível de significância 5%, ou seja, não há efeito das drogas. Em outras palavras, não há evidências estatísticas suficientes para comprovar a existência de efeito das drogas ao nível de significância de 5%.

4.1.3 Análise de Efeito de Período

Para o cálculo do efeito de período são usadas as médias da Tabela 10. O estimador não-viciado para o efeito de período P pode ser descrito como:

$$\begin{aligned} \hat{P} &= \frac{1}{2} \left[(\bar{Y}_{.21} - \bar{Y}_{.11}) - (\bar{Y}_{.12} - \bar{Y}_{.22}) \right] = [(401,8292 - 405,0017) - (512,2117 - 455,5467)]/2 = \\ &= -29,91875. \end{aligned}$$

O cálculo da estatística de teste é dado por

$$T_p = \frac{\hat{P}}{\hat{\sigma}_d \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = -29,91875/17,24416467 = -1,735,$$

que ao ser comparada com o valor tabelado t = 2,074, leva à aceitação da hipótese nula. Ou seja, pode-se dizer que não há evidências estatísticas suficientes para comprovar a existência de efeito de período ao nível de significância de 5%.

4.1.4 Análise de Variância

Os resultados da Análise de Variância para a área sobre a curva de concentração do tempo zero ao tempo t estão sintetizados na Tabela 13. Os cálculos foram computados a partir das expressões apresentadas na seção 3.2.4, considerando o modelo completo.

Tabela 13. Análise de Variância para ASC_t com o modelo completo.

Fonte de Variação	GL	SQ	QM	F	Valor p
Entre Indivíduos					
Efeito Residual (seqüência)	1	77.692,98	77.692,98	1,62	0,22
Resíduo (entre indivíduos)	22	1.056.882,06	48.040,09		
Dentro de Indivíduos					
Droga (tratamento)	1	8.571,51	8.571,51	2,40	0,14
Período	1	10.727,22	10.727,22	3,00	0,10
Resíduo (dentro de indivíduos)	22	78.501,55	3.568,25		
Total	47	1.232.375,32			

Os resultados indicam que o efeito residual não é significativo (p -valor = 0,22), podendo então o modelo crossover ser reduzido, onde nesse novo modelo a fonte de variação resíduo (entre indivíduos) terá seu efeito incorporado ao efeito residual (seqüência), já que o mesmo não é significativo, aumentando-se os graus de liberdade.

Os resultados dessa análise de variância com o modelo reduzido encontra-se na Tabela 14.

Tabela 14. Análise de Variância para *ASCt* com o modelo reduzido.

Fonte de Variação	GL	SQ	QM	F	Valor p
Entre Indivíduos					
Sujeitos (indivíduos)	23	1.134.575,04	49.329,35	13,82	0,0001
Dentro de Indivíduos					
Droga (tratamento)	1	8.571,51	8.571,51	2,40	0,14
Período	1	10.727,22	10.727,22	3,00	0,10
Resíduo (dentro de indivíduos)	22	78.501,55	3.568,25		
Total	47	1.232.375,32			

Nota-se que o efeito entre indivíduos é significativo a um nível de significância de 5%, o que indica a presença de variabilidade entre os indivíduos. Os efeitos de período e de droga não são significativos a 5%. As conclusões, evidentemente foram as mesmas que as obtidas com o teste t.

A análise de variância deve ser realizada com base no modelo reduzido.

4.1.5 Intervalos de Confiança para Bioequivalência Média

Considerar-se-á os intervalos de confiança para a bioequivalência média para a medida *ASCt* com e sem transformação, ou seja, na escala original e logarítmica, respectivamente. Na prática é necessário apresentar a análise de bioequivalência em apenas uma das escalas, de acordo com o resultado das análises descritiva e de adequabilidade do modelo do planejamento crossover.

4.1.5.1 Intervalos de Bioequivalência

Para a construção dos intervalos necessita-se a média da variável *ASCt* para o tratamento de referência. Da Tabela 8 têm-se $\bar{y}_R = 430,2742$. Então os limites para o intervalo de confiança para a diferença das médias são dados por:

$$\hat{\theta}_{\text{inf}} = -0,20\bar{y}_R = -0,20*430,2742 = -86,0548$$

$$\hat{\theta}_{\text{sup}} = 0,20\bar{y}_R = 0,20*430,2742 = 86,0548.$$

Os intervalos para a diferença das médias na escala logarítmica e os intervalos de confiança para a razão das médias, tanto na escala original quanto logarítmica são fixos, não dependendo dos resultados da amostra. Os intervalos de confiança encontram-se na Tabela 15.

Tabela 15. Intervalos de Confiança para a diferença das médias.

Escala	Diferença $\theta_{\text{inf}} < \mu_T - \mu_R < \theta_{\text{sup}}$	Razão $\delta_{\text{inf}} < \mu_T / \mu_R < \delta_{\text{sup}}$
Original	(-86,0548 ; 86,0548)	(80% ; 120%)
Logarítmica	(-0,2231 ; 0,2231)	(80% ; 125%)

4.1.5.2 Intervalo de Confiança para a Bioequivalência: Método Paramétrico

Os cálculos necessários para se aplicar os métodos paramétricos baseados no intervalo de bioequivalência para as médias e razões das formulações R e T de *ASCt* foram retirados das Tabelas 8 e 9.

Os limites calculados para o IC 95% de confiança para a diferença das médias para os dados sem transformação (escala original) são:

$$\text{LIC: } (\bar{y}_T - \bar{y}_R) - t_{0,05/2}(22) * \text{E.P} = (457,0204 - 430,2742) - 2,074 * 17,2440 = -9,0176$$

$$\text{LSC: } (\bar{y}_T - \bar{y}_R) + t_{0,05/2}(22) * \text{E.P} = (457,0204 - 430,2742) + 2,074 * 17,2440 = 62,51$$

Os limites calculados para o IC 95% de confiança para a razão das médias para os dados sem transformação (escala original) são:

$$\text{LIC: } [(-9,0176/430,2742) + 1] * 100\% = 97,9042\%$$

$$\text{LSC: } [(62,5100/430,2742)+1]*100\% = 114,5279\%$$

E para os dados transformados, os limites para o intervalo com 95% de confiança para a diferença das médias são calculados do mesmo modo que na escala original, mas com as medidas farmacocinéticas transformadas em logaritmo natural, e são dados por:

$$\text{LIC: } (5,9914 - 6,0640) - 2,074 * 0,06304 = -0,2033$$

$$\text{LSC: } (5,9914 - 6,0640) + 2,074 * 0,06304 = 0,0581$$

A partir desses limites, é calculado o intervalo para a razão das médias com 95% de confiança. Os limites de confiança são:

$$\text{LIC: } [\exp(-0,2033)] * 100\% = 81,6033\%$$

$$\text{LSC: } [\exp(0,0581)] * 100\% = 105,9821\%$$

Os resultados encontram-se sumarizados na Tabela 16.

Tabela 16. Intervalos de Confiança.

Escala	IC para a diferença das médias	IC para a razão das médias
Original	(-9,0176 ; 62,51)	(97,9042% ; 114,5279%)
Logarítmica	(-0,2033 ; 0,0581)	(81,6033% ; 105,9821%)

Segundo os critérios utilizados para definir os intervalos de bioequivalência paramétricos, e com base nos limites especificados, pode-se ver que todos os intervalos da Tabela 16 estão contidos nos respectivos intervalos de bioequivalência da Tabela 15. E a partir disso, conclui-se pela bioequivalência média das formulações R e T, sob a análise da medida $ASCt$.

4.1.5.3 Intervalo de Confiança para a Bioequivalência: Método Não-Paramétrico

Como sugere a literatura, o intervalo não-paramétrico é calculado usando-se 90% de confiança. O intervalo de 90% de confiança não-paramétrico baseado no estimador de Hodges-Lehmann é construído utilizando as $12 \times 12 = 144$ diferenças das diferenças de períodos entre as seqüências 1 e 2 ordenadas. Os limites do intervalo de 95% de confiança para $\theta = \mu_T - \mu_R$ foram

encontrados por $LD_{inf} = D(w(5\%)) = D(43) = -8,675$ e por $LD_{sup} = D(w(95\%)+1) = D(102) = 58,84$. Portanto, como o intervalo $(-8,675 ; 58,84)$ está contido em $(-86,0548 ; 86,0548)$, a bioequivalência entre R e T é verificada, considerando a análise de ASC_{t_K} .

4.2 Análise Preliminar do Exemplo usando variável Cmax

As análises serão realizadas com os dados referentes às concentrações plasmáticas ao longo do tempo das Tabelas 6 e 7 e com a variável Cmax (concentração máxima observada), que se encontra nas Tabelas 8 e 9.

4.2.1 Análise Descritiva

A avaliação estatística da bioequivalência deve ser iniciada por uma análise descritiva dos dados, tanto para as observações originais (medidas de concentração tomadas ao longo do tempo) como para as medidas farmacocinéticas. Além da familiarização com os dados, esta análise preliminar auxilia na verificação de suposições necessárias para a aplicação dos testes estatísticos propostos, bem como por possibilitar uma tomada de decisão mais fidedigna nas análises subsequentes para determinar se há bioequivalência entre as formulações comparadas (R e T).

4.2.1.1 Análise Descritiva das Concentrações Plasmáticas

A análise descritiva das concentrações plasmáticas nos intervalos de tempo definidos permite visualizar as diferenças entre as formulações R e T utilizando os dados originais obtidos através de um experimento crossover 2x2.

A Tabela 17 apresenta as estatísticas descritivas para as concentrações plasmáticas referentes ao medicamento de referência aplicado nos 24 indivíduos (sujeitos) e a Tabela 18 refere-se ao medicamento de teste aplicado nas mesmas condições.

Tabela 17. Concentrações Plasmáticas (ng/ml) ao longo do tempo (horas) do medicamento de referência.

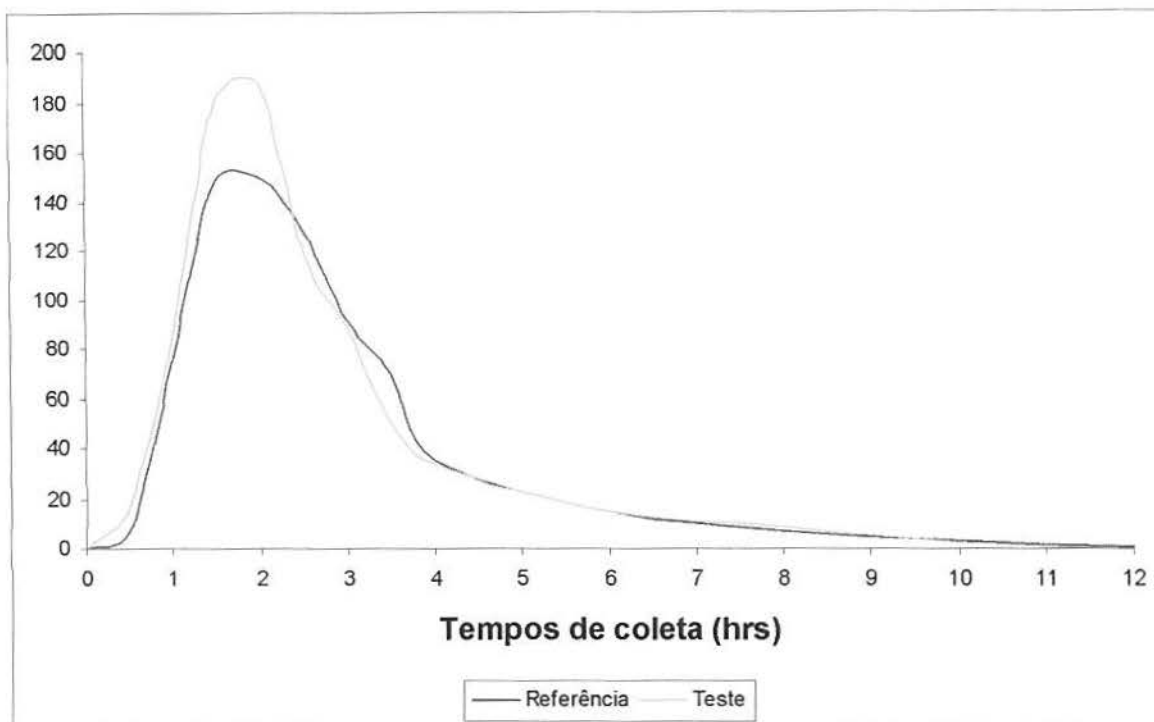
Tempo	Média	Mediana	Desvio-padrão	Erro-padrão	CV(%)	Mínimo	Máximo
0	0	0	0	0	-	0	0
0,5	6,99	0	10,85	2,22	155,33	0	32,5
1	75,45	43,25	95,16	19,42	126,12	0	357,1
1,5	148,15	115,2	116,04	24,2	78,32	0	515,5
2	149,86	146,3	75,61	15,43	50,45	23,5	317,2
2,5	128,12	117,95	75,42	15,39	58,86	25,1	297,4
3	91,08	65,9	56,46	11,52	61,99	14,23	217,9
3,5	69,93	56,9	47,28	9,65	67,6	11,5	177,9
4	35,24	29,3	24,62	5,03	69,88	0	81,3
6	15,04	13,7	13,26	2,71	88,14	0	37,2
8	7,34	0	8,79	1,79	119,7	0	22,8
10	2,87	0	5,86	1,2	204,56	0	17,5
12	0,69	0	3,37	0,69	489,9	0	16,5

Tabela 18. Concentrações Plasmáticas (ng/ml) ao longo do tempo (horas) do medicamento de teste.

Tempo	Média	Mediana	Desvio-padrão	Erro-padrão	CV(%)	Mínimo	Máximo
0	0	0	0	0	-	0	0
0,5	17,24	0	29,76	6,07	172,59	0	129,7
1	85,81	56,3	87,92	17,95	102,46	0	291,4
1,5	180,17	188,4	118,08	24,1	65,54	12,4	586,2
2	187,76	200,25	87,47	17,85	46,59	38,9	335,5
2,5	118,49	93,4	67,03	13,68	56,57	22,5	297,7
3	87,66	63,9	68,87	14,06	78,57	23,7	324,5
3,5	49,29	43,4	27,46	5,61	55,71	11,3	116,9
4	33,42	27,8	19,61	4,0	58,67	10,7	88,2
6	15,26	15,35	11,58	2,36	75,87	0	38,2
8	8,39	8,95	8,7	1,77	103,67	0	23,2
10	2,36	0	5,47	1,12	231,37	0	17,1
12	0	0	0	0	-	0	0

Para a visualização do comportamento das concentrações ao longo do tempo, que refletem a absorção, distribuição e eliminação dos medicamentos R e T recomenda-se utilizar o gráfico das médias das concentrações em função do tempo (Gráfico 1). Também torna-se interessante adicionar a esse gráfico a variação de ± 2 erros-padrões.

Gráfico 1. Concentração média (ng/ml).

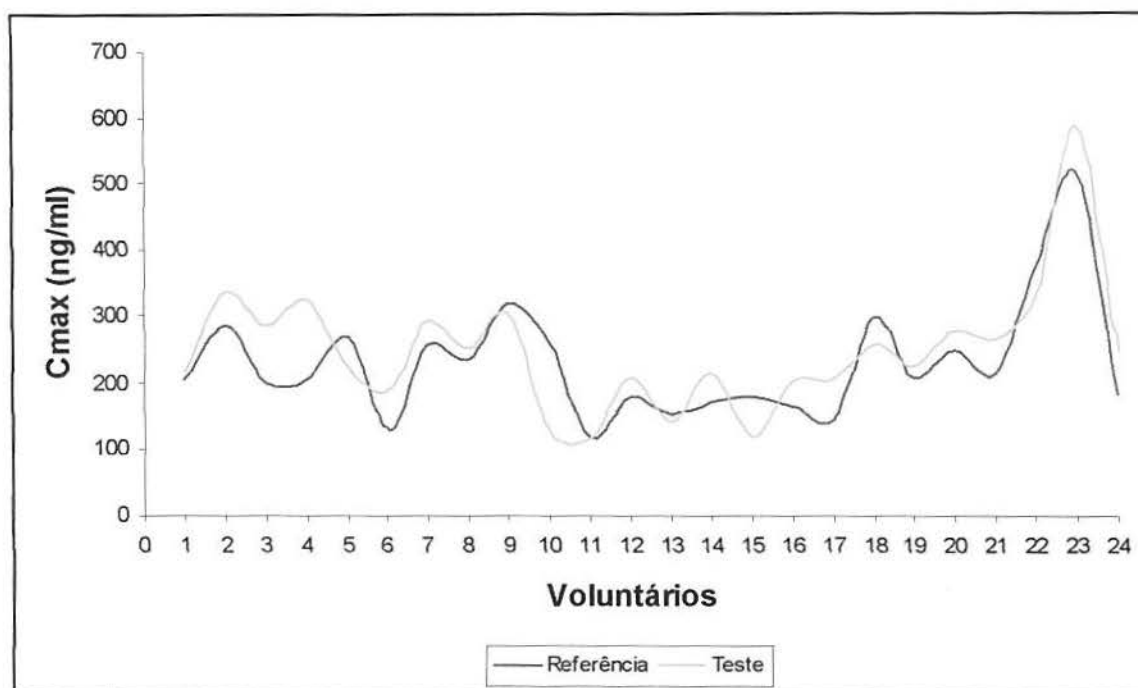


Pode-se observar pelo gráfico que o medicamento teste apresenta concentração média superior em praticamente todos os pontos em relação ao medicamento de referência. À medida que o tempo aumenta, essa diferença diminui, tornando-se quase nula, pelo que se percebe pelas linhas quase sobrepostas. Entre os períodos 1 e 2,5, a concentração plasmática média para ambos medicamentos tem um considerável aumento, e após o tempo 4 de coleta, a concentração diminui exponencialmente.

Com base nas Tabelas 17 e 18, pode-se observar que as médias, medianas e desvios-padrões crescem e depois decrescem ao longo do tempo, refletindo o comportamento natural do fármaco no organismo referente às fases de absorção até a eliminação dos mesmos. Nota-se também grande variabilidade das medidas de concentração, especialmente na vizinhança de C_{max} , ou seja, nos tempos correspondentes às maiores concentrações e altos valores para os coeficientes de variação, o que significa uma grande heterogeneidade dos indivíduos (sujeitos) quanto às concentrações plasmáticas. O Gráfico 2 apresenta as concentrações máximas para cada voluntário do experimento, mostrando as tendências das médias das concentrações. Percebe-se que a concentração máxima possui um comportamento semelhante para as formulações de

referência e teste, embora haja para a grande maioria dos voluntários uma superioridade na concentração máxima para a formulação de teste.

Gráfico 2. Concentração Máxima para cada voluntário.



4.2.1.2 Análise Descritiva das Medidas Farmacocinéticas

A análise descritiva prossegue para as medidas farmacocinéticas correspondentes às concentrações plasmáticas individuais e às concentrações médias dos sujeitos. Essas medidas resumem os dados multivariados (concentrações plasmáticas repetidas ao longo do tempo) em observações univariadas, as quais refletem o comportamento de absorção e eliminação dos fármacos R e T.

As estatísticas para as concentrações individuais (para cada fármaco) estão nas Tabelas 19 e 20.

Tabela 19. Estatísticas Descritivas das Medidas Farmacocinéticas para o Medicamento de Referência.

Medida Farmacocinética	Média	Mediana	Erro-padrão	CV(%)	Mínimo	Máximo
Tmax	1,92	2,00	0,12	31,41	1,00	3,50
Cmax	229,73	207,55	17,78	37,91	117,60	515,50
K_e	0,57	0,45	0,07	63,90	0,08	1,21
t1/2	2,01	1,56	0,36	88,24	0,57	8,66
ASC _t	430,27	397,78	33,36	37,98	175,25	765,05
ASC _∞	466,18	416,14	37,68	39,59	182,85	798,33

Tabela 20. Estatísticas Descritivas das Medidas Farmacocinéticas para o Medicamento de Teste.

Medida Farmacocinética	Média	Mediana	Erro-padrão	CV(%)	Mínimo	Máximo
Tmax	1,75	2,00	0,10	27,94	1,00	3,00
Cmax	247,72	238,55	19,68	38,92	117,00	586,20
K_e	0,44	0,38	0,06	68,00	0,07	1,19
t1/2	2,53	1,84	0,41	80,13	0,58	9,55
ASC _t	457,01	449,14	33,23	35,62	225,38	844,33
ASC _∞	495,21	518,33	36,01	35,62	233,71	926,39

Observa-se um mesmo padrão das medidas farmacocinéticas para os dois medicamentos (R e T) que também é semelhante com o da média. Os valores dos coeficientes de variação diminuíram sensivelmente comparados com os CV's das concentrações plasmáticas.

4.3 Análise de Variância das Medidas Farmacocinéticas

Para a análise de variância das medidas farmacocinéticas serão usados os dados das Tabelas 8 e 9.

4.3.1 Análise de Variância das Medidas Farmacocinéticas para o Modelo Completo

As análises foram realizadas utilizando a opção modelos lineares do módulo Anova do Analyst do SAS, onde as variáveis classificatórias são os sujeitos, o período, a seqüência e a droga. O efeito de seqüência será testado com o termo de erro sujeito(seqüência), devido à sua

peculiar estrutura. As variáveis dependentes serão Tmax (tempo onde foi alcançada a concentração máxima), Cmax (concentração máxima observada), Ke (constante de velocidade de eliminação total do fármaco no organismo), t1/2 (meia-vida de eliminação do fármaco), ASCt (área sobre a curva de concentração do tempo 0 ao tempo t) e ASCinf (área sobre a curva de concentração do tempo 0 ao tempo infinito).

4.3.1.1 Análise de Variância para a variável Tmax

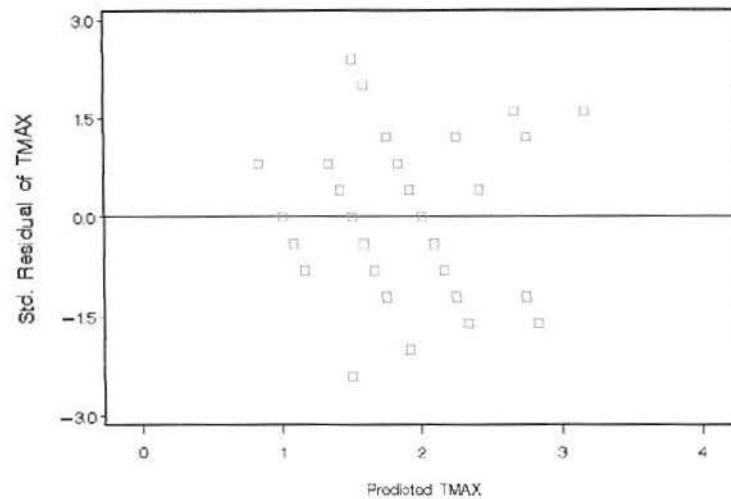
Através da Tabela 21, pode-se concluir que o efeito de sujeito (seqüência) é significativo a 5%, ou seja, há diferença significativa entre os sujeitos para cada seqüência com relação ao tempo onde foi alcançada a concentração máxima. Também pode-se perceber que os efeitos de seqüência (efeito residual), droga e período não foram significativos a 5%.

Tabela 21. Teste dos efeitos fixos do modelo usando a variável Tmax.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	0,3333	0,3333	0,66	0,4247
Sujeito(seqüência)	22	11,0833	0,5038	5,32	0,0001
Droga	1	0,3333	0,3333	3,52	0,0740
Período	1	0,3333	0,3333	3,52	0,0740

Através do gráfico dos resíduos contra os valores preditos pelo modelo (Gráfico 3), nota-se que não há nenhum indício de estrutura de funil, o que sugere a satisfação da homogeneidade de variância.

Gráfico 3. Valores preditos x Resíduos para a variável Tmax.



Para a verificação da normalidade foram realizados os testes de Anderson-Darling e Kolmogorov-Smirnov, que produziram os resultados apresentados na Tabela 22.

Tabela 22. Testes de normalidade para a variável Tmax.

Teste	Estatística	Valor p
A-D	0,6356	0,094
K-S	0,1458	0,012

Os resultados da Tabela 22 e o Gráfico 4 sustentam a suposição de normalidade. No Gráfico 5 não há nenhum indicador de outliers.

Gráfico 4. P-P-Plot dos resíduos para a variável Tmax.

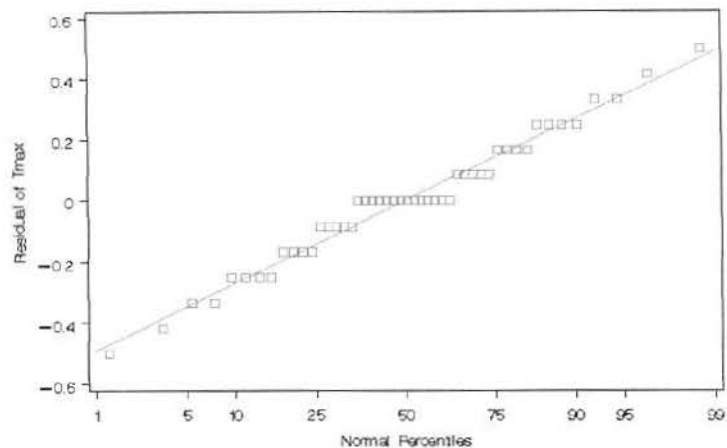
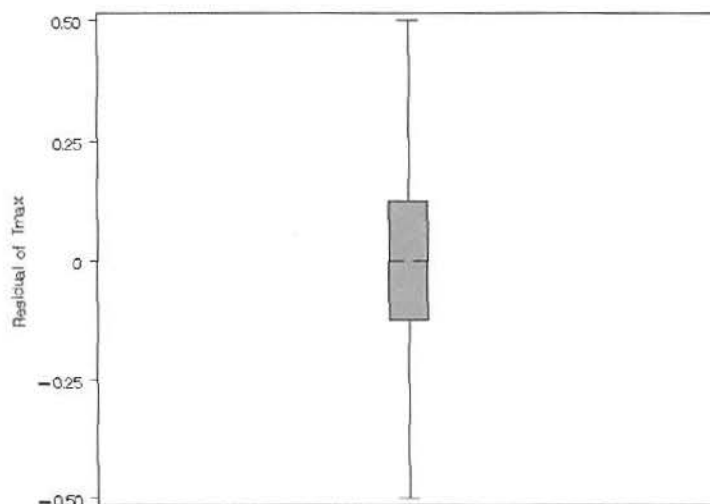


Gráfico 5. Boxplot dos resíduos do modelo para a variável Tmax.



4.3.1.2 Análise de Variância para a variável Cmax

A Tabela 23 sintetiza os resultados dos testes, e nota-se que o efeito de sujeito (seqüência) apresenta significância estatística. Em outras palavras, há diferença significativa entre os sujeitos para cada seqüência com relação à concentração máxima observada. Também pode-se perceber que os efeitos de seqüência (efeito residual), droga e período não foram significativos a 5%.

Tabela 23. Teste dos efeitos fixos do modelo usando a variável Cmax.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	233,64187	233,64187	0,01	0,9050
Sujeito(seqüência)	22	352.831,5379	16.037,7972	10,37	<0,0001
Droga	1	3.886,2002	3.886,2002	2,51	0,1272
Período	1	1.167,2269	1.167,2269	0,75	0,3944

Os Gráficos 6, 7 e 8 e a Tabela 24 mostram que as suposições do modelo não foram violadas. No Gráfico 6 há dois valores discrepantes dos demais, mas nada que comprometa a análise. O Gráfico 7 mostra uma boa aderência à normalidade, devido à grande concentração dos valores em torno da reta ajustada, fato também comprovado pelos testes de normalidade da Tabela 24. E o Gráfico 8 mostra a dispersão dos valores preditos e dos resíduos do modelo ajustado para a variável Cmax, sugerindo homocedasticidade dos resíduos.

Tabela 24. Testes de normalidade para a variável Cmax.

Teste	Estatística	Valor p
A-D	0,18383	>0,25
K-S	0,06620	>0,15

Gráfico 6. Boxplot dos resíduos para a variável Cmax.

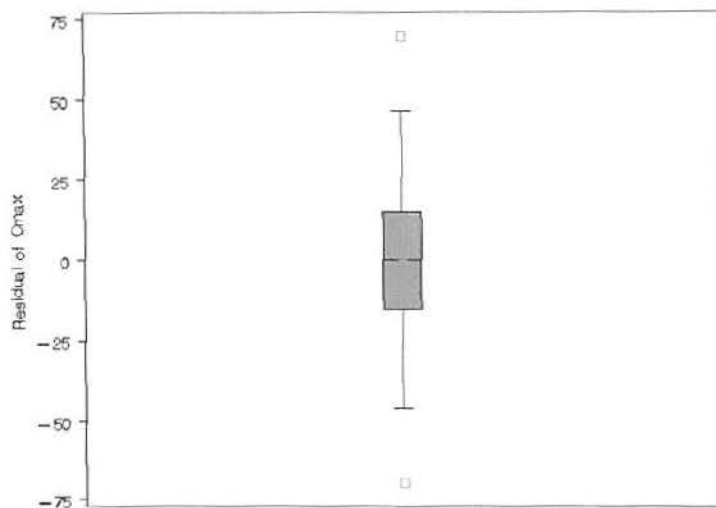


Gráfico 7. P-P-Plot dos resíduos para a variável Cmax.

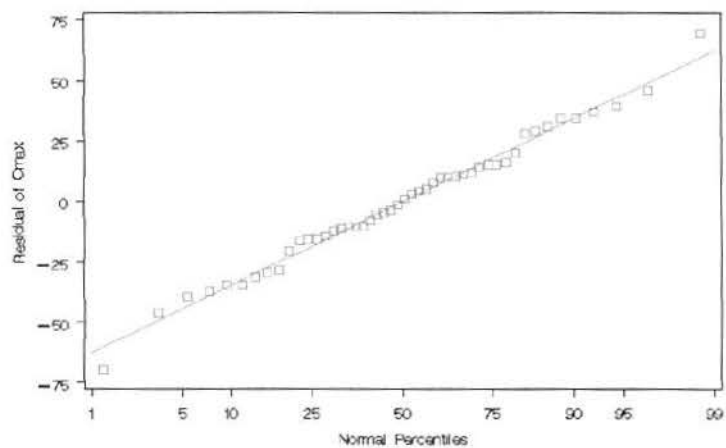
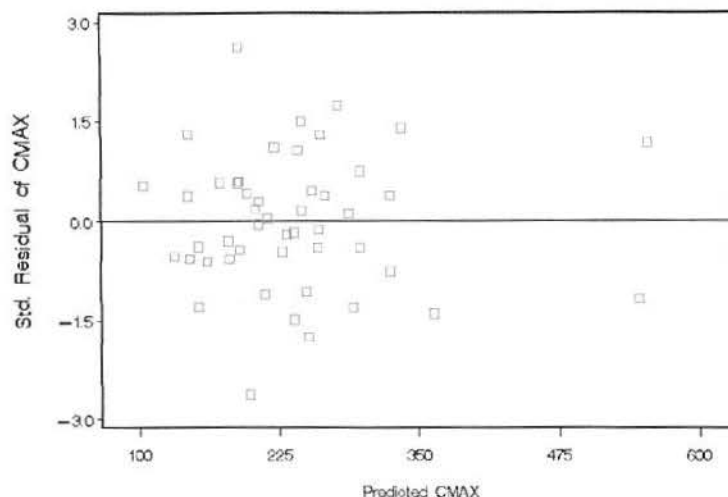


Gráfico 8. Valores preditos x Resíduos para a variável Cmax.



4.3.1.3 Análise de Variância para a variável Ke

A Tabela 25 sintetiza os resultados dos testes, e nota-se que além do efeito de sujeito(seqüência), a droga também apresenta significância estatística. Em outras palavras, há diferença significativa entre as formulações de referência e teste quanto à constante de velocidade de eliminação total do fármaco no organismo a 5%. E também, há diferença significativa entre os sujeitos para cada seqüência com relação à constante de velocidade de eliminação total do fármaco no organismo. Também pode-se perceber que os efeitos de seqüência (efeito residual) e período não foram significativos a 5%.

Tabela 25. Teste dos efeitos fixos do modelo usando a variável Ke.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	0,03413	0,03413	0,18	0,6763
Sujeito(seqüência)	22	4,19376	0,19062	5,02	0,0002
Droga	1	0,21333	0,21333	5,62	0,0270
Período	1	0,00800	0,00800	0,21	0,6506

Analisando a Tabela 26 vê-se que a suposição de normalidade está garantida. Ao se analisar os gráficos, pode-se perceber que não há nenhum indicador de violação das suposições da análise.

Tabela 26. Testes de normalidade para a variável Ke.

Teste	Estatística	Valor p
A-D	0,3943	>0,25
K-S	0,0892	>0,15

Gráfico 9. Valores preditos x Resíduos para a variável Ke.

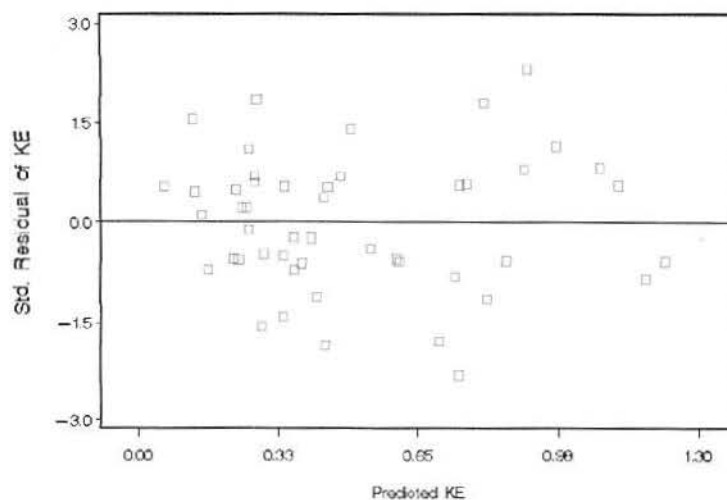


Gráfico 10. Boxplot dos resíduos para a variável Ke.

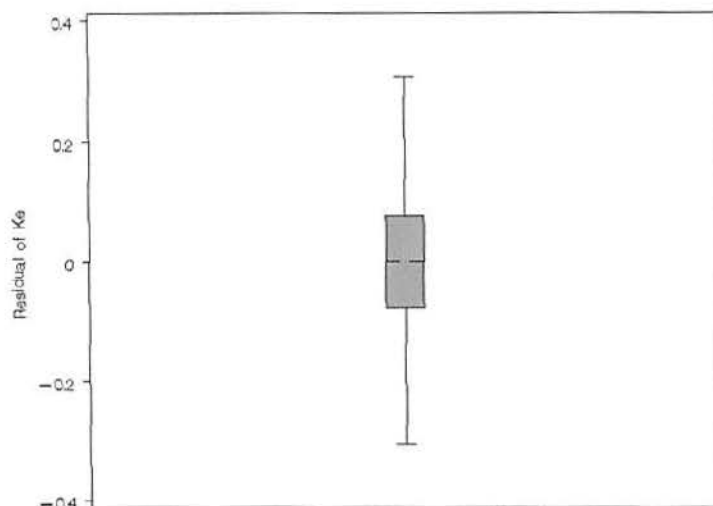
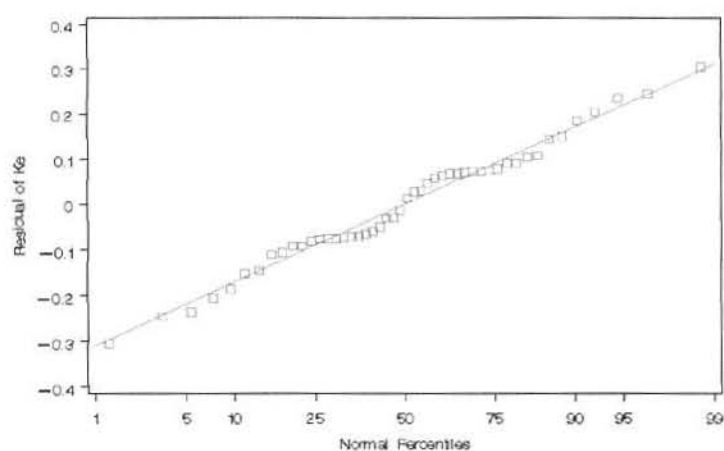


Gráfico 11. P-P-Plot dos resíduos para a variável Ke.



4.3.1.4 Análise de Variância para a variável t1/2

A Tabela 27 sintetiza os resultados dos testes, e nota-se que não há nenhum efeito significativo, mostrando que não há diferenças significativas entre os sujeitos dentro de cada

seqüência quanto à meia-vida de eliminação do fármaco a 5%. O mesmo acontece com os efeitos de seqüência, droga e período.

Tabela 27. Teste dos efeitos fixos do modelo usando a variável t1/2.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	0,6348	0,6348	0,14	0,7140
Sujeito(seqüência)	22	101,3273	0,1906	1,69	0,1129
Droga	1	3,1519	3,1519	1,16	0,2937
Período	1	4,9152	4,9152	1,80	0,1929

Analisando a tabela 28 vê-se que a suposição de normalidade não está satisfeita, fato também comprovado pelo gráfico de probabilidade normal, onde a não aderência à reta ajustada é evidente. E ao analisar os demais gráficos pode-se perceber que há uma estrutura de funil na dispersão dos resíduos e valores discrepantes no boxplot, evidenciando a necessidade de alguma transformação nos dados, a fim de atender às exigências do modelo.

Tabela 28. Testes de normalidade para a variável t1/2.

Teste	Estatística	Valor p
A-D	1,9467	<0,005
K-S	0,1570	<0,01

Gráfico 12. Valores preditos x Resíduos para a variável t1/2.

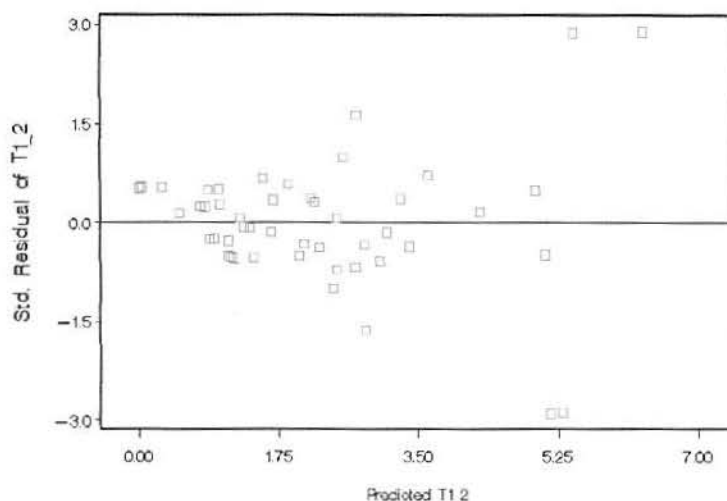


Gráfico 13. Boxplot dos resíduos para a variável t1/2.

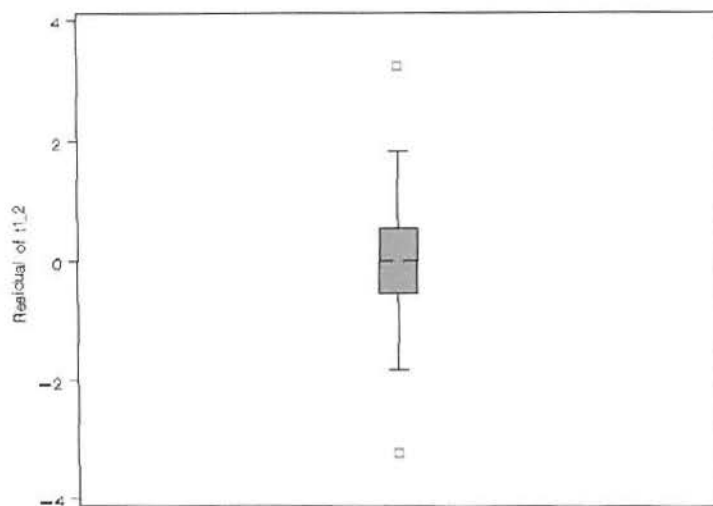
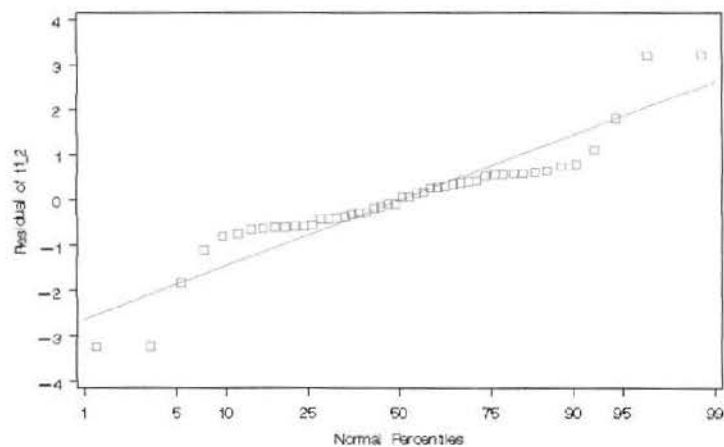


Gráfico 14. P-P-Plot dos resíduos para a variável t1/2.



4.4.1.5 Análise de Variância para a variável ASCt

A Tabela 29 sintetiza os resultados dos testes, e nota-se que há apenas um efeito significativo, mostrando que há diferenças significativas entre os sujeitos dentro de cada seqüência quanto à área sobre a curva de concentração do fármaco do tempo zero ao tempo t a

um nível de significância de 5%. Os efeitos de seqüência, droga e período não foram significativos a 5%.

Tabela 29. Teste dos efeitos fixos do modelo usando variável ASCt.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	77.692,9807	77.692,9807	1,62	0,2168
Sujeito(seqüência)	22	1.056.900,7380	48.040,943	13,46	<0,0001
Droga	1	8.584,3430	8.584,3430	2,41	0,1352
Período	1	10.741,5790	10.741,5790	3,01	0,0967

Analisando a Tabela 30 vê-se que a suposição de normalidade está satisfeita, fato também comprovado pelo gráfico de probabilidade normal (Gráfico 17), onde a maioria dos valores ficaram próximos à reta ajustada. E ao analisar os demais gráficos (15 e 16) pode-se perceber que não há nenhuma estrutura de funil na dispersão dos resíduos ou valores discrepantes no boxplot, o que indica adequabilidade do modelo.

Tabela 30. Teste de normalidade para a variável ASCt.

Teste	Estatística	Valor p
A-D	0,1298	>0,25
K-S	0,0526	>0,15

Gráfico 15. Valores preditos x Resíduos para a variável ASCt.

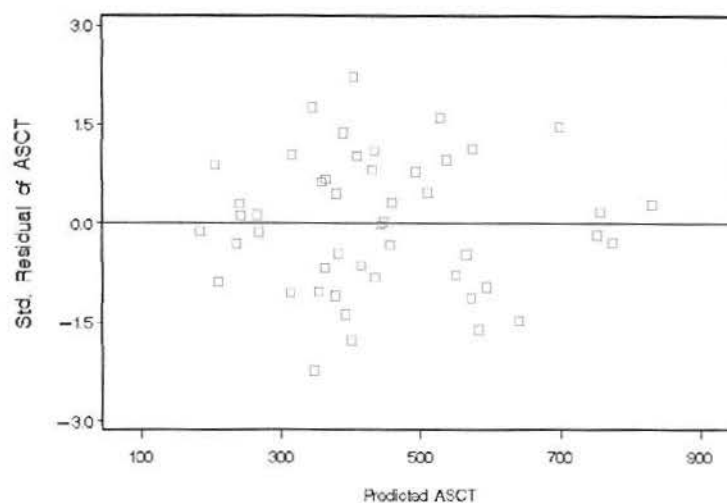


Gráfico 16. Boxplot dos resíduos para a variável ASCt.

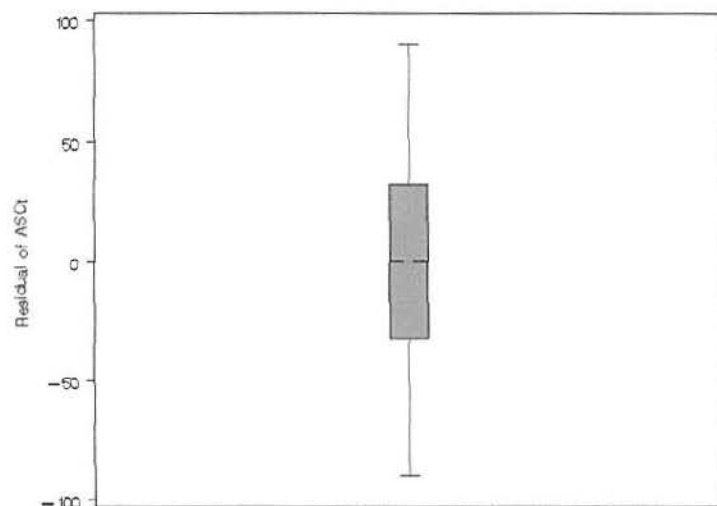
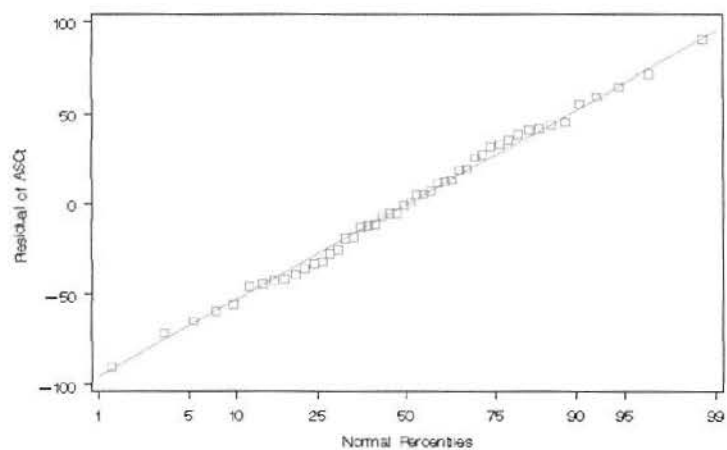


Gráfico 17. P-P-Plot dos resíduos para a variável ASCt.



4.3.1.6 Análise de Variância para a variável ASCinf

A tabela 31 sintetiza os resultados dos testes, e nota-se que há efeito significativo de sujeito(seqüência) e de período, mostrando que há diferenças significativas entre os sujeitos dentro de cada seqüência quanto à área sobre a curva de concentração do fármaco do tempo zero ao tempo infinito e que há também diferença significativa entre os períodos quanto à área sobre a

curva de concentração do fármaco do tempo zero ao tempo infinito a um nível de significância de 5%. Os efeitos de seqüência e droga não foram significativos a 5%.

Tabela 31. Teste dos efeitos fixos do modelo usando a variável ASCinf.

Efeito	GL	SQ	QM	F	Valor p
Seqüência	1	59.463,36047	59.463,36047	1,00	0,3275
Sujeito(seqüência)	22	1.304.313,826	59.286,992	11,55	<0,0001
Droga	1	10.103,313	10.103,313	1,97	0,1746
Período	1	22.696,126	22.696,126	4,42	0,0472

Analisando a Tabela 32 vê-se que a suposição de normalidade está satisfeita, dado os altos valores de probabilidade, fato também comprovado pelo gráfico de probabilidade normal (Gráfico 20), onde a maioria dos valores ficaram próximos à reta ajustada. E ao analisar os demais gráficos (Gráficos 18 e 19) pode-se perceber que não há nenhuma estrutura de funil na dispersão dos resíduos ou valores discrepantes no boxplot, o que indica adequabilidade do modelo.

Tabela 32. Testes de normalidade para a variável ASCinf.

Teste	Estatística	Valor p
A-D	0,2652	>0,25
K-S	0,0691	>0,15

Gráfico 18. Valores preditos x Resíduos para a variável ASCinf.

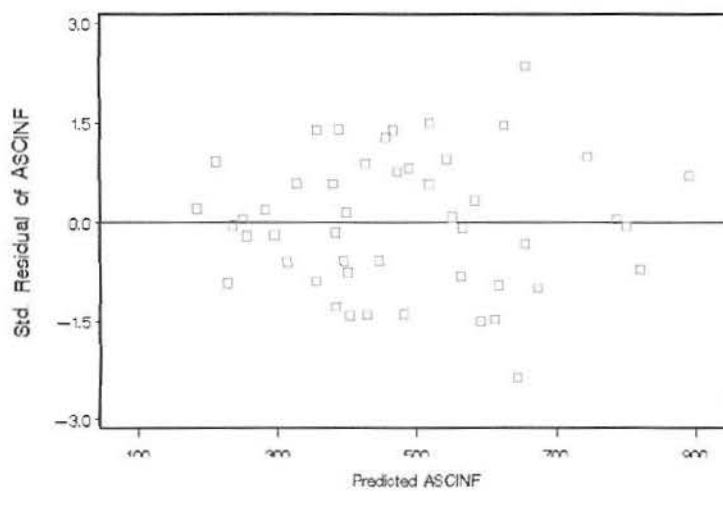


Gráfico 19. Boxplot dos resíduos para a variável ASCinf.

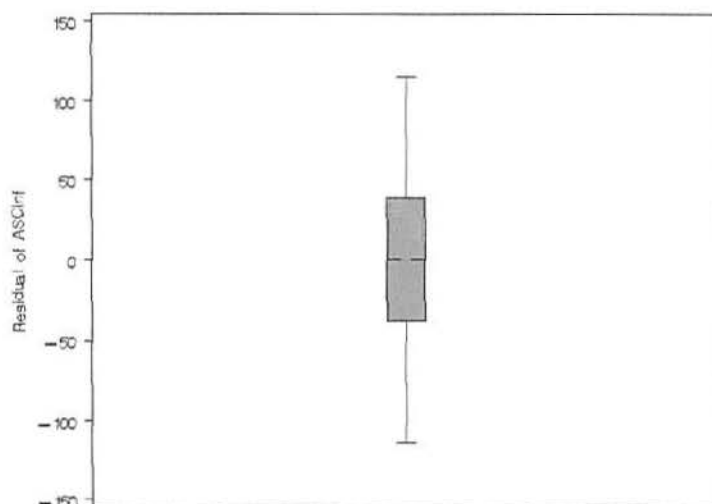
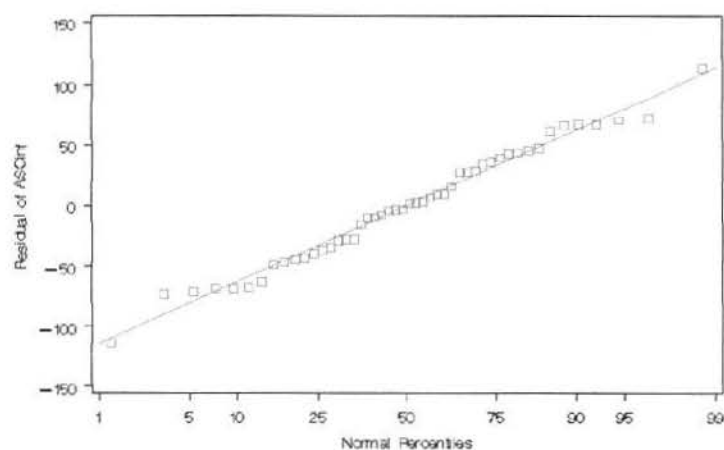


Gráfico 20. P-P-Plot dos resíduos para a variável ASCinf.



Como visto nas análises anteriores, a única variável que apresentou problemas de satisfação das suposições foi a variável meia-vida $t_{1/2}$. Foram refeitas as análises com a variável transformada em logaritmo natural e raiz quadrada, mas não houve melhorias importantes na adequação. Com as transformações contemplou-se a normalidade, muito embora a permanência de variância não constante e presença de outliers.

4.3.2 Análise de Variância das Medidas Farmacocinéticas para o Modelo Reduzido

As análises realizadas com o modelo reduzido estão sintetizadas na Tabela 33, onde foi retirado o efeito de seqüência do modelo e incorporado ao efeito de sujeitos. Os resultados e as interpretações são as mesmas que em 4.3.1.

Tabela 33. Teste dos efeitos fixos do modelo usando as medidas farmacocinéticas.

Significância	Tmax	Cmax	Ke	t1/2	ASCt	ASCinf
Sujeitos	0,0001	<0,0001	0,0002	0,1289	<0,0001	<0,0001
Droga	0,0740	0,1272	0,0270	0,2937	0,1352	0,1746
Período	0,0740	0,3944	0,6506	0,1929	0,0967	0,0472

4.4 Análise de Medidas Repetidas da Variável Concentração Plasmática ao longo do Tempo

Para a análise de medidas repetidas da concentração plasmática foram usados os dados das Tabelas 6 e 7. A análise foi realizada no software SAS versão 8.2, usando a opção Medidas Repetidas do procedimento Anova do Analyst. A variável dependente é a concentração plasmática, e as variáveis classificatórias são os sujeitos, a seqüência, o período, a droga (tratamento) e o tempo de coleta. Foram retirados os dados referentes ao tempo zero, pois só havia valores nulos, não se tornando informativos. O efeito de sujeito é considerado dentro de seqüência, isto é, sujeito (seqüência). Serão incorporadas ao modelo as interações simples entre o tempo e droga e tempo e período. A estrutura de covariância adotada foi a Componentes Simétricos, que além de ser a mais restritiva e exigente, também possui o menor número de parâmetros a serem estimados, o que garante uma convergência para as soluções das estimativas dos parâmetros, fato que não ocorreu para as demais estruturas disponíveis na opção Medidas Repetidas do SAS com qualquer um dos três métodos de estimação dos parâmetros de covariância (REML, ML e MIVQUE0). A Tabela 34 mostra os resultados para os testes para os efeitos fixos do modelo completo.

Tabela 34. Testes para os efeitos fixos do modelo completo.

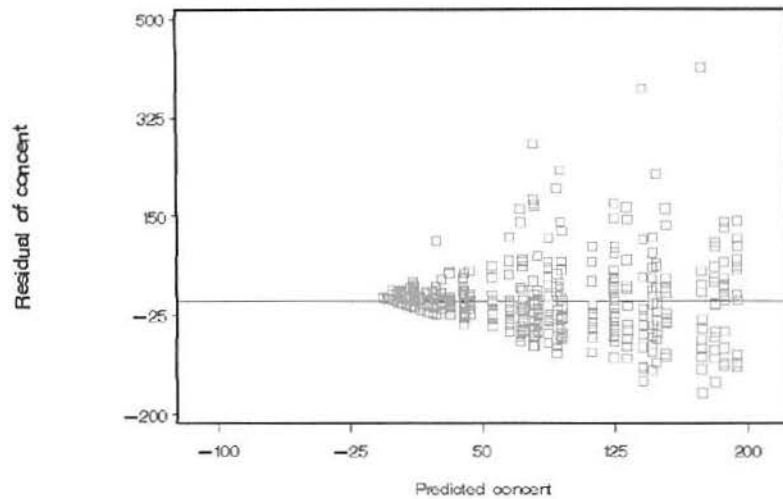
Efeito	GL numerador	GL denominador	F	Valor p
Seqüência	1	22	1,60	0,2197
Período	1	517	0,36	0,5486
Tempo	11	517	59,75	<0,0001
Droga	1	517	1,17	0,2801
Tempo*Droga	11	517	1,11	0,3479
Tempo* Período	11	517	0,45	0,9307

Com base na Tabela 34, pode-se observar que o único efeito significativo foi o efeito de tempo, ou seja, há diferença significativa quanto às concentrações plasmáticas entre os diferentes tempos de coleta, a um nível de significância de 5%. Os demais efeitos de seqüência, período, droga, tempo*droga e tempo*período não foram significativos a 5%, mostrando que não há diferença significativa quanto às concentrações plasmáticas entre as seqüências, os períodos e as drogas. Também pode-se dizer que não há evidências estatísticas suficientes para concluir quanto à existência de efeito conjunto de droga e tempo e de tempo e período em relação à concentração plasmática, a um nível de significância de 5%.

A verificação de adequação do modelo foi realizada com base nos gráficos dos resíduos (Gráficos 21 a 23).

No Gráfico 21 é nítida uma dispersão em forma de funil, indicando assim falha na suposição de homocedasticidade ou variância constante.

Gráfico 21. Valores Preditos x Resíduos para modelo completo.



Pelo Gráfico 22 pode-se ver vários valores discrepantes, o que novamente sugere uma violação nas suposições do modelo. O gráfico de probabilidade normal (Gráfico 23) e os testes de aderência à normalidade mostram que essa exigência também não foi satisfeita. (Anderson-Darling: $p < 0,005$ e K-S: $p < 0,01$).

Gráfico 22. Boxplot dos resíduos para modelo completo.

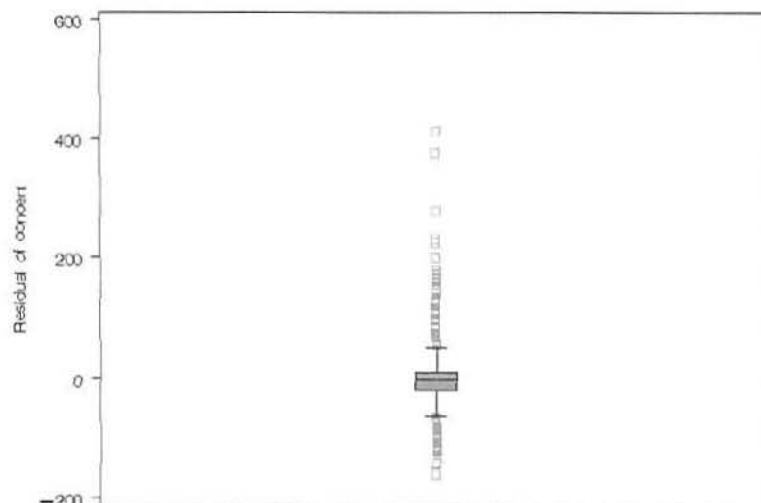
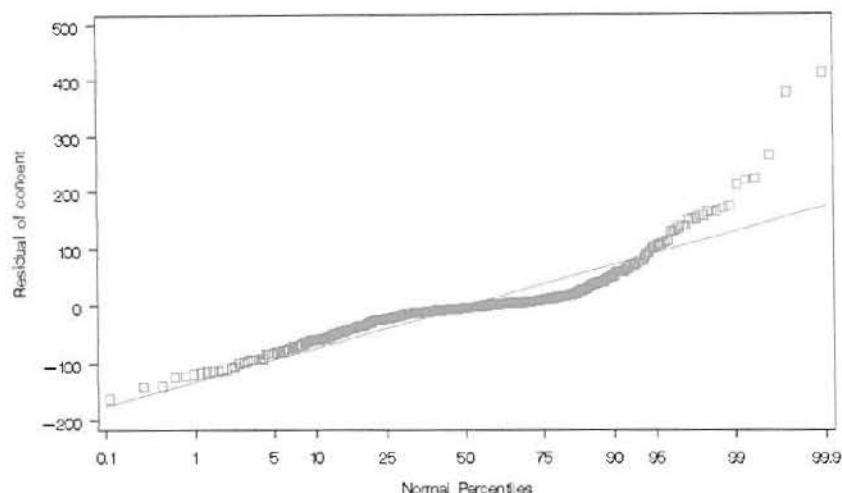


Gráfico 23. P-P-Plot dos resíduos para modelo completo.



Os Gráficos 24, 25, 26 e 27 apresentam a dispersão dos resíduos em função, respectivamente, do tempo, droga, seqüência e período. Nota-se que há grande variabilidade dos resíduos em função do tempo. Pode-se perceber também que nos tempos mais baixos (até 4 horas) a variabilidade é bem maior, e à medida que o tempo passa, essa variabilidade diminui, conforme o Gráfico 24. Para os demais efeitos (droga, seqüência e período) os resíduos apresentam comportamento semelhante, de acordo com os Gráficos 25 a 27. E a partir, disso, pode-se concluir que a variabilidade que está comprometendo o bom ajuste do modelo está presente no tempo, fazendo com que seja ajustada uma regressão em função do tempo, para assim, ver qual a transformação necessária. Foi ajustada uma regressão linear simples entre o logaritmo do desvio-padrão com o logaritmo da média para os níveis de tempo, obtendo-se um coeficiente de regressão linear de aproximadamente 0,5 o que indica uma transformação raiz quadrada aos dados. Dada a existência de valores nulos, adicionou-se a constante 1 aos dados antes de se determinar a raiz. Com os dados transformados, procedeu-se a análise e obteve-se os resultados da Tabela 35.

Gráfico 24. Resíduos em função do tempo.

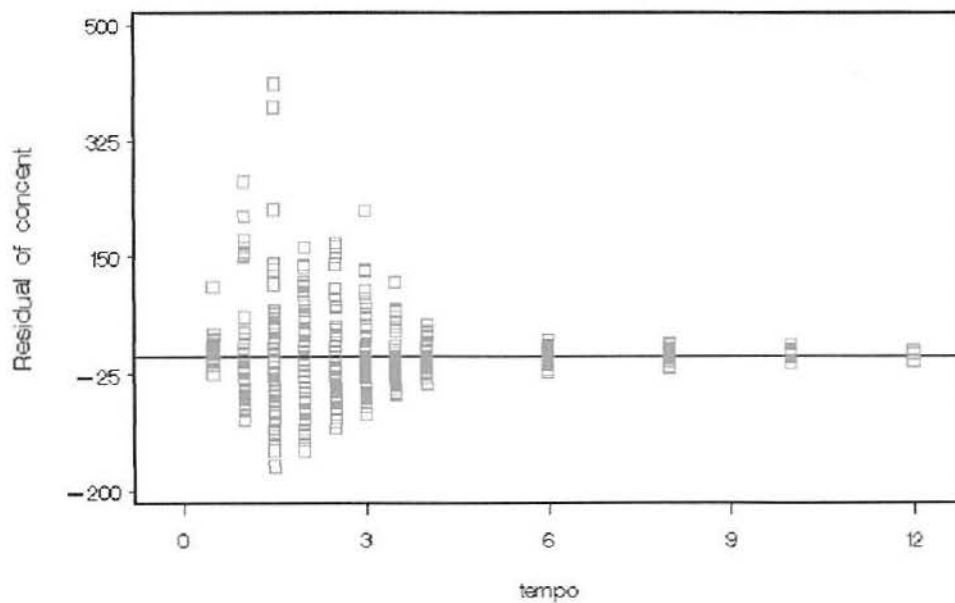


Gráfico 25. Resíduos em função das drogas (formulações).

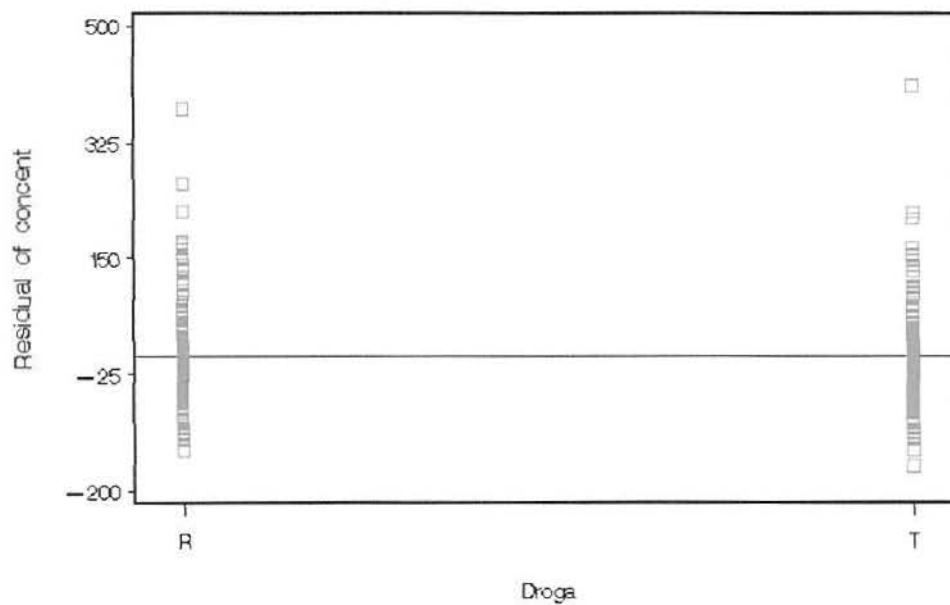


Gráfico 26. Resíduos em função das seqüências.

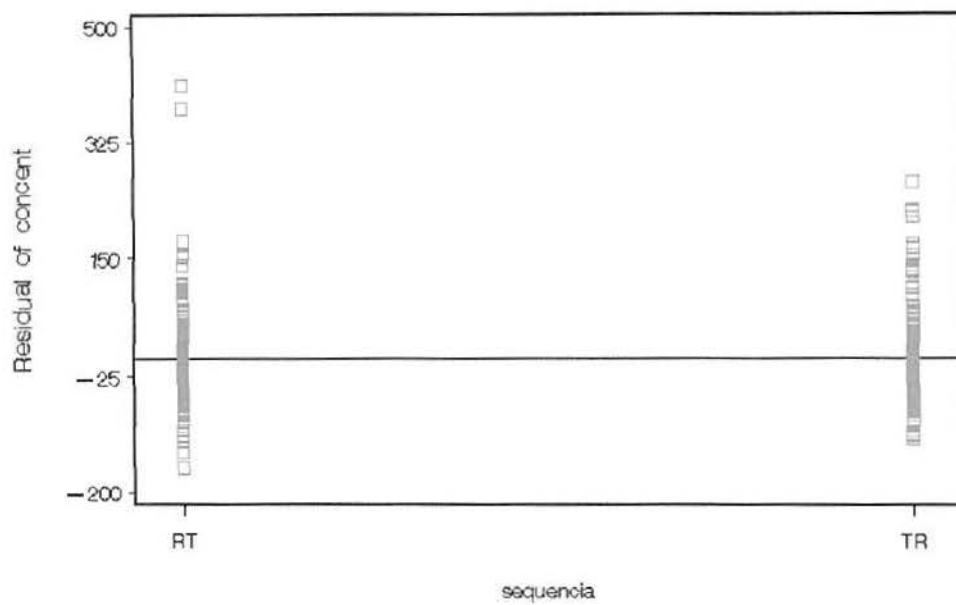
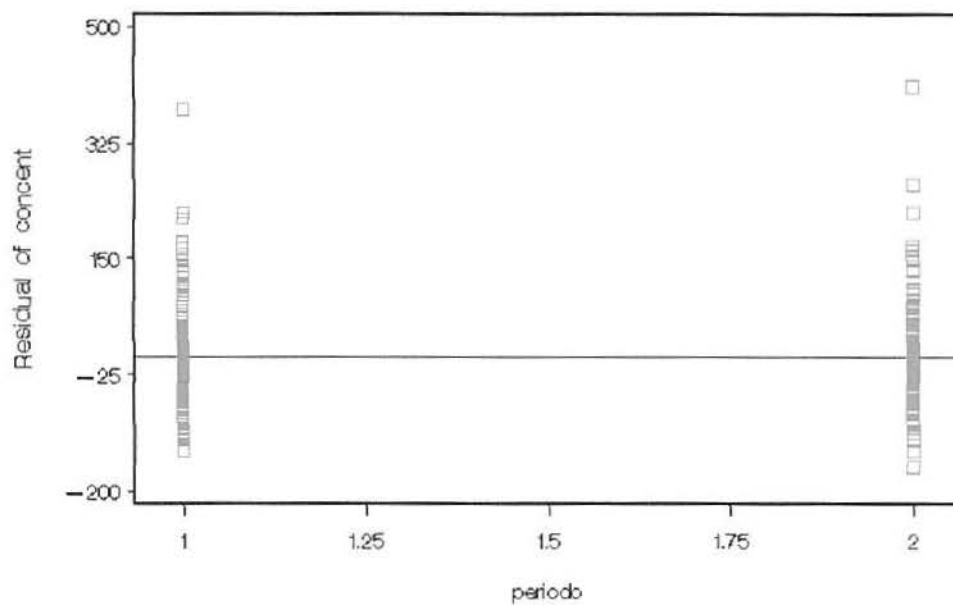


Gráfico 27. Resíduos em função dos períodos.



A Tabela 35 apresenta os testes para os efeitos fixos para os dados transformados. Os resultados em termos de significância dos efeitos não se alteraram.

Tabela 35. Testes para os efeitos fixos do modelo com os dados transformados.

Efeito	GL numerador	GL denominador	F	Valor p
Seqüência	1	22	2,07	0,1643
Período	1	517	0,30	0,5836
Tempo	11	517	114,93	<0,0001
Droga	1	517	1,71	0,1913
Tempo*Droga	11	517	1,07	0,3803
Tempo* Período	11	517	0,60	0,8316

Pelos gráficos percebe-se alguma melhoria em relação à adequação do modelo. No Gráfico 28 dos valores preditos contra os resíduos a estrutura de funil está aparentemente amenizada. O boxplot apresenta alguma melhoria e o gráfico de probabilidade normal também mostra uma melhoria sensível dos pontos dispersos no ajuste à reta, indicando que a suposição de normalidade está satisfeita.

Gráfico 28. Valores preditos x Resíduos para modelo completo com os dados transformados.

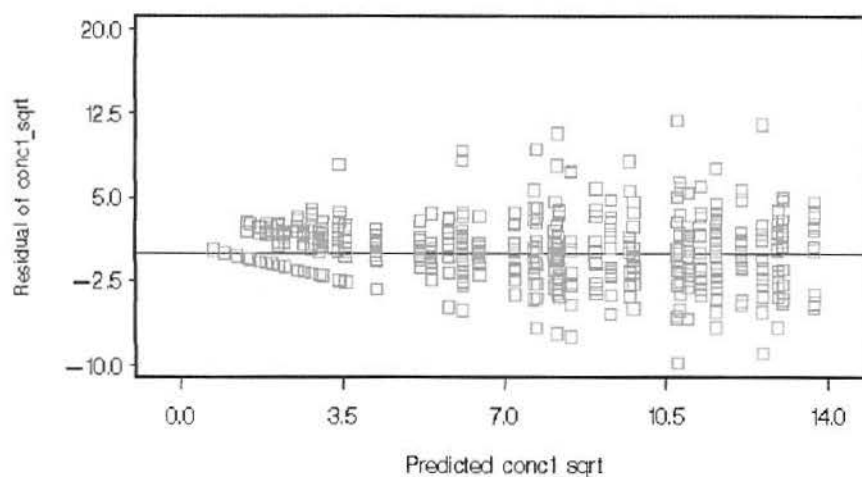


Gráfico 29. Boxplot dos resíduos para modelo completo com a variável transformada.

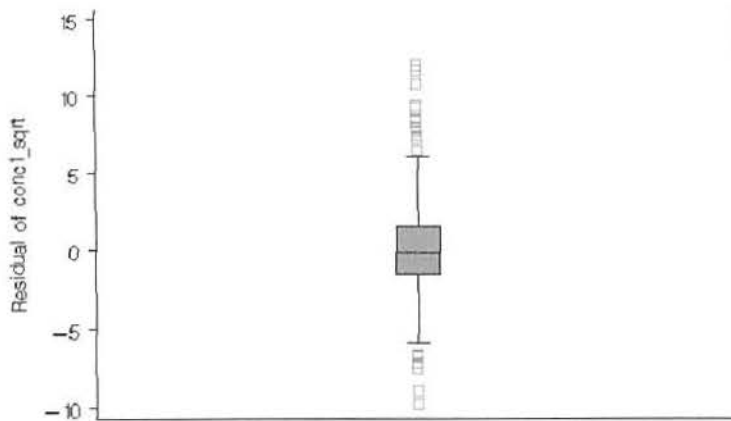
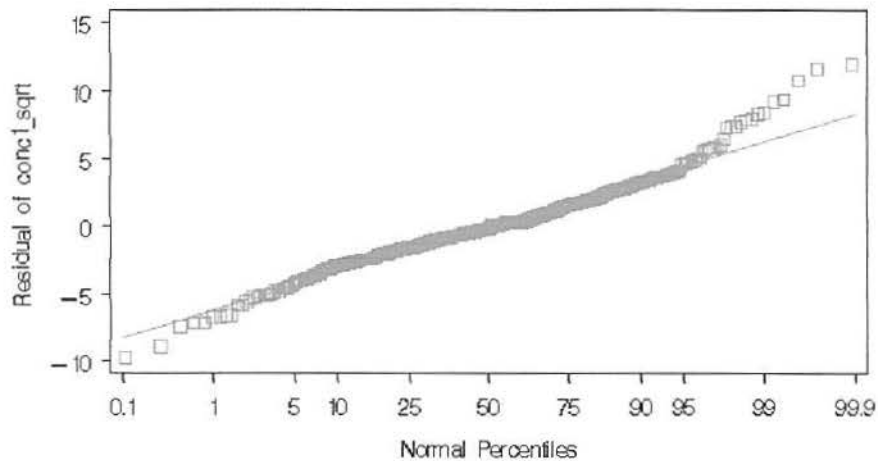


Gráfico 30. P-P-Plot dos resíduos para modelo completo com a variável transformada.



Recomenda-se também realizar as análises usando um modelo reduzido, já que a seqüência não teve efeito significativo. No modelo reduzido, os efeitos fixos serão o período, a droga, o tempo e as interações simples entre o tempo e droga e o tempo e período. A única diferença entre os modelos completo e reduzido é a retirada do efeito de seqüência, ou melhor, o efeito da seqüência será incorporado ao efeito residual entre os indivíduos (sujeitos). A Tabela 36 mostra os testes para os efeitos fixos do modelo.

Tabela 36. Testes para os efeitos fixos do modelo reduzido para a concentração plasmática transformada.

Efeito	GL numerador	GL denominador	F	Valor p
Período	1	517	0,30	0,5836
Tempo	11	517	114,43	<0,0001
Droga	1	517	1,71	0,1913
Tempo*Droga	11	517	1,07	0,3803
Tempo* Período	11	517	0,60	0,8316

Com base na Tabela 36, pode-se ver que o único efeito significativo novamente foi o efeito de tempo, mostrando que há diferenças significativas quanto às concentrações plasmáticas entre os diferentes tempos de coleta, a um nível de significância de 5%. Os efeitos restantes não foram significativos a 5%.

Ao proceder-se a verificação das suposições do modelo, percebe-se que há um comportamento muito semelhante ao ajuste do modelo completo com a variável concentração plasmática transformada, satisfazendo assim as exigências do modelo.

5. PROCEDIMENTOS COMPUTACIONAIS DE ANÁLISE

Neste capítulo serão apresentados alguns tópicos sobre softwares estatísticos que analisam dados referentes a planejamentos crossover característicos de ensaios de bioequivalência.

A maioria dos pacotes estatísticos calcula apenas o tamanho amostral ou o poder desejado para a existência de bioequivalência, como, por exemplo, o SAS, que fornece o cálculo do tamanho de amostra para três opções: One Sample Equivalence, Paired Equivalence e Two Sample Equivalence. A partir da opção escolhida de modelo (aditivo ou multiplicativo), o programa calcula tanto o tamanho de amostra quanto o poder. Para essas análises são necessárias as especificações de teste, como a diferença nula, o coeficiente de variação, o nível de significância e o poder. Também são calculados os limites inferior e superior de bioequivalência. A Figura 1 mostra uma das janelas do SAS, onde foi escolhida a opção Sample Size, One Sample Equivalence, e especificados os valores necessários (diferença nula = 0,10 e CV de 30%). Os limites de bioequivalência escolhidos foram -0,20 e +0,20.

Figura 1. One Sample Equivalence.

One: Enter a Command X			
Calculate:		<input type="radio"/> Power	<input checked="" type="radio"/> N
Model:		<input checked="" type="radio"/> Additive	<input type="radio"/> Multiplicative
Test specifications			
Null difference:	0.10		
Coeff of variation:	0.30		
Alpha:	0.05		
Power:	From:	To:	By:
	0.90		
Plot		Equivalence bounds	
<input checked="" type="checkbox"/> Power vs. N			
Power ref line:		Lower: -0.20	
N ref line:		Upper: +0.20	
<input type="button" value="OK"/> <input type="button" value="Cancel"/> <input type="button" value="Reset"/> <input type="button" value="Save Options"/> <input type="button" value="Help"/> <input type="button" value="Title"/>			

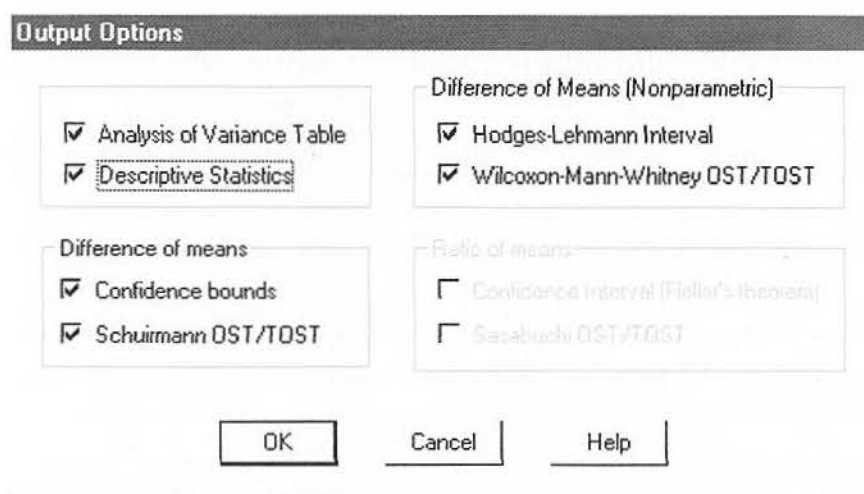
Os resultados produziram uma amostra de tamanho 79 com poder de 90%. Para as outras duas opções, os procedimentos são os mesmos.

O SAS também realiza outras formas de análises com procedimentos gerais sem usar técnicas específicas para bioequivalência, mas que podem ser usadas para tal, como por exemplo, análise de variância, análise de medidas repetidas, entre outros. Os resultados computacionais apresentados no Capítulo 4 foram produzidos através do SAS.

Um dos softwares estatísticos específicos para testes de bioequivalência é o EquivTest, um dos mais recentes softwares liberados pela Statistical Solutions Ltda e aprovado pelo FDA (Food and Drug Administration) e CPMP (Committee for Proprietary Medicinal Products), além do nQuery, pertencente ao mesmo grupo.

O EquivTest, disponível em www.statsol.ie/equivtest/equivtest.htm, além de ser de fácil uso computacional, também apresenta saídas (outputs) claras, completas e editáveis. O programa calcula as medidas farmacocinéticas facilmente através de programações ou transformações. Outra grande qualidade do programa em questão é a alta capacidade de armazenamento dos dados e a elasticidade para aceitar diversas estruturas para as variáveis. No EquivTest podem ser calculados os testes e intervalos de confiança paramétricos e não-paramétricos ao mesmo tempo, especificando a escolha de teste uni ou bilateral, como mostra a Figura 2.

Figura 2. Testes e Intervalos de Confiança.



Os usuários também podem definir os parâmetros de bioequivalência, optando por diferença ou razão de médias ou proporções ou aproveitar as opções padrões pré-programadas, como mostram as Figuras 3 e 4.

Figura 3. Especificações do teste de equivalência.

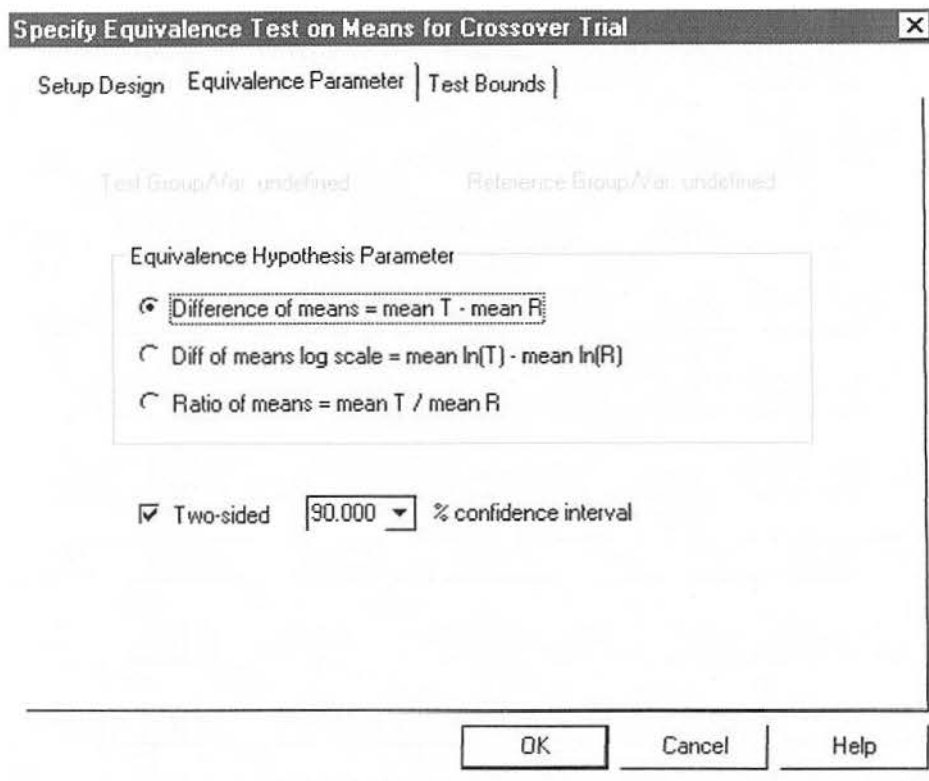


Figura 4. Especificações do teste de equivalência.

Specify Equivalence Test on Means for Crossover Trial

Setup Design | Equivalence Parameter | Test Bounds |

Parameter: $\text{Difference of means} = \text{mean T} - \text{mean R}$

Test

- Lower bound only
- Upper bound only
- Symmetric upper and lower alpha values
- Asymmetric

Equivalence Bounds

Alpha value	Lower	Upper	Upper alpha value
0.0500	-0.20MeanF	0.20MeanR	

OK Cancel Help

Outro programa estatístico usado para a análise de ensaios de bioequivalência é o nQuery Advisor (www.statsol.ie/nquery/review3.htm), que por ser da mesma companhia que o EquivTest, apresenta estrutura semelhante, não sendo necessário repetir os mesmos detalhes. A Figura 5 mostra uma das janelas do programa.

Figura 5. Janela do programa nQuery.

nQuery Advisor - [MTE4-1] - □ ×

File Edit View Options Assistants Plot Window Help - | 5 | ×

Crossover design TOST for ratio of means (using original scale)			
	1	2	3
Test significance levels, α (one-sided)	0.050	0.050	
Lower equivalence limit for $\mu_T / \mu_S, \Delta_L$	0.900	0.850	
Upper equivalence limit for $\mu_T / \mu_S, \Delta_U$	1.111	1.176	
Expected ratio, μ_T / μ_S	1.000	1.000	
Coefficient of variation between subjects, CV_B	0.206	0.206	
Coefficient of variation, intrasubject, CV_I	0.158	0.158	
Power (%)	90	90	
n per group	26	12	

Crossover design TOST for ratio of means (using original scale)

Enter a value for alpha, α , the significance level for the t-test. Enter the lower equivalence limit for the ratio of means, the upper equivalence limit, the expected true ratio of means, and the coefficients of variation for between- and within-subject variation, and then specify power or sample size and nQuery Advisor will compute the other. Data are assumed to be normally distributed in the original scale. hit

For Help, press F1 0.05000000 AUTO RECALC OFF NUM

6. CONSIDERAÇÕES FINAIS

Os Ensaio de Bioequivalência vem ganhando cada vez mais reconhecimento e importância, embora ainda sejam recentes no Brasil, pois passaram a ser utilizados de forma rotineira apenas a partir da Lei nº 9787/99 dos genéricos. Tais estudos são de inigualável valor, pois testam novas drogas, que poderão substituir os remédios convencionais, muito caros para a maioria dos brasileiros. Segundo a revista americana *The Economist*, o Brasil é o 9º país do mundo em consumo de medicamentos per capita, mas 50% dos pacientes que precisam de um medicamento não podem comprá-lo e abandonam o tratamento, salientando mais uma vez a utilidade da realização de estudos de biodisponibilidade e bioequivalência para a utilização de medicamentos genéricos.

Os Ensaio de Bioequivalência têm uma série de normas e diretrizes a serem cumpridas, desde a fase de planejamento experimental até às análises estatísticas, que devem satisfazer integralmente os critérios exigidos pela Resolução da ANVISA ANVS/MS391/99.

Neste trabalho procurou-se aprimorar a linguagem estatística adotada pelo Manual da ANVISA, dando-lhe mais consistência e rigor estatístico. Acrescentou-se também os procedimentos bootstrap, tema muito focado no contexto de bioequivalência em trabalhos científicos apresentados nos congressos e simpósios estatísticos, e a análise de medidas repetidas, que não foram contempladas pela ANVISA, mas merecem atenção especial.

Outro elemento muito importante é a caracterização do planejamento experimental para tais estudos. A ANVISA recomenda o modelo crossover. Deve-se usar delineamentos que identifiquem e isolem as fontes de variabilidade entre os indivíduos na análise dos dados. Após a escolha de um planejamento experimental adequado, outros aspectos merecem uma devida atenção, como a aleatorização, para que sejam válidas as inferências estatísticas, os efeitos residuais, o período de eliminação “washout” e o número de voluntários. A ocorrência de efeitos residuais compromete toda análise, sendo sugerido pela ANVISA, que a análise seja feita apenas com base nos dados referentes ao primeiro período. Quanto à escolha do tamanho amostral, o protocolo da ANVISA recomenda um mínimo de 24 sujeitos, muito embora haja aumento do poder com o acréscimo do tamanho de amostra.

Outro aspecto interessante de ser considerado é a escala utilizada na análise dos dados. O Manual da ANVISA induz o leitor a transformar as variáveis (medidas farmacocinéticas) em

logaritmo natural para então proceder as análises. Conceitualmente a transformação preventiva não é correta, apenas transforma-se a variável quando seu comportamento na escala original não satisfaz as suposições necessárias de independência, homogeneidade de variâncias e normalidade.

A ANVISA também apresenta um modelo estatístico para a análise de variância próprio, mas podem ser incluídas a esse modelo interações e efeitos que o estatístico julgue ser relevante, fazendo com que o modelo fique mais completo e detalhado. O mesmo pode ser dito para a análise de medidas repetidas para dados em planejamento crossover usando a metodologia de modelos mistos.

A realização deste trabalho evidenciou a importância do profissional de Estatística na área de ensaios de bioequivalência. A ausência de um profissional da área para a execução das análises estatísticas, além de outros aspectos, como a falta de integração entre as etapas de estudo, dificuldades com relação à execução correta do delineamento experimental, desistências dos voluntários no decorrer do estudo, falta de conhecimento das técnicas por parte do responsável pela execução da etapa e a falta de um estatístico no planejamento dos estudos, evidencia-se a necessidade da participação direta do profissional de Estatística na área da saúde, principalmente no que diz respeito à pesquisa clínica, evitando desperdício de tempo e dinheiro.

Acredita-se que os objetivos deste trabalho tenham sido alcançados. Alguns aspectos, como os procedimentos bootstrap e a análise de medidas repetidas, não foram suficientemente aproveitados, devido à amplitude do assunto. Muitos dos tópicos abordados neste trabalho poderiam ser estudados de forma mais completa, propiciando temas de futuros trabalhos na área.

O Capítulo 4 é seguramente de grande importância para o trabalho, dada a descrição e análise de um exemplo de um ensaio de bioequivalência. Acredita-se que com a sua leitura é possível ter uma boa noção de como lidar com esse tipo de informação, uma vez que apresenta com riqueza de detalhes os procedimentos e interpretações necessárias.

Em resumo, acredita-se que este trabalho possa ser um bom referencial, tanto teórico quanto aplicado, àqueles que iniciarem os estudos nos ensaios de bioequivalência, tema extremamente importante e útil.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ALBANESE, M.T., (1990). *Latent Variable Models for Binary Response Data*. Ph.D. Dissertation. The London School of Economics and Political Science, ENG.
- AMBROSANO, G.M.B. & CORRENTE, J.E., (2003). Intervalos de Confiança Bootstrap em Ensaio de Bioequivalência, **10º SEAGRO**, UFLA, Lavras-MG, 7 a 11 de julho de 2003.
- ANVISA (Agência Nacional de Vigilância Sanitária), Manual de Boas Práticas em Biodisponibilidade e Bioequivalência. (www.anvisa.gov.br)
- CHOW, S.C. & LIU, J.P., (1992). *Design and Analysis of Bioavailability and Bioequivalence Studies*. New York, Marcel Decker.
- CHOW, S.C. & LIU, J.P., (2000). *Design and Analysis of Bioavailability and Bioequivalence Studies*. New York, Marcel Decker.
- EFRON, B., (1979). Bootstrap methods: another look at the jackknife. **Annals of Statistics**.
- EFRON, B., GONG, G., (1983). A leisurely look at the bootstrap, the jackknife and cross validation. **Annals of Statistics**, 37: 36-45.
- EquivTest (Software Estatístico) www.statsol.ie/equivtest/equivtest.htm
- FLEISS, J.L., (1999). *Design and Analysis of Clinical Experiments*. New York, John Wiley & Sons, Inc.
- GRIEVE, A.P., (1982). The two-period changeover design in clinical trials. **Biometrics** 38, 517.
- GRIZZLE, J.E., (1965). The two-period changeover design and its use in clinical trials. **Biometrics** 25, 357-381.

HILLS, M. & ARMITAGE, P., (1979). The two-period cross-over clinical trial. **British Journal of Clinical Pharmacology** 8, 7-20.

JOHN, P.W.M., (1971). *Statistical Design and Analysis of Experiments*. New York, Macmillan.

JONES, B. & KENWARD. M.G., (1990). *Design and Analysis of Cross-Over Trials*. Second Edition. New York, Chapman & Hall.

MANLY, B.F.J., (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Second Edition. New York, Chapman & Hall.

MIRANDA, C.G. & SIQUEIRA, A.L., (2003). Inferência Estatística para Estudos de Bioequivalência Média, 10º SEAGRO, UFLA, Lavras-MG, 7 a 11 de julho de 2003.

MONTGOMERY, D.C., (2000). *Design and Analysis of Experiments*. Fifth Edition. New York, Wiley & sons, Inc.

NQuery (Software Estatístico) www.statsol.ie/nquery/review3.htm

SATTERHWAITE, F.E., (1946). An approximate distribution of estimates of variance components. **Biometrics**, 110–112.

SCHUIRMANN, D.J., (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. **Journal of Pharmacokinetics and Biopharmaceutics** 15, 657-680.

SCHUIRMANN, D.J., (1989). Treatment of bioequivalence data: log transformation. **Proceedings of Bio-International 89 - Issues in the Evaluation of Bioavailability Data**, Toronto, Canada, October 1-4, 159-61.

TUKEY, J.W., (1958). The teaching of concrete mathematics. **American Mathematical Monthly** 65 (January),1-9.

WESTLAKE, W.J., (1973), *The Design and Analysis of Comparative Blood-Level Trials*. Ed. J. Swarbrick.

WESTLAKE, W.J., (1988), *Bioavailability and Bioequivalence of Pharmaceutical Formulations*. K.E. Peace, Marcel Dekker, Inc.

WADA, C.Y., (2003). Métodos Estatísticos para Ensaio de Bioequivalência em Produtos Farmacêuticos, 10º SEAGRO, UFLA, Lavras-MG, 7 a 11 de julho de 2003.

8. ANEXOS: GLOSSÁRIO

1. Analito: Composto químico específico. Pode ser um fármaco intacto, biomolécula ou seu derivado, metabólito ou produto de degradação em uma matriz biológica.
2. Biodisponibilidade Relativa: Quociente da quantidade e velocidade de princípio ativo que chega à circulação sistêmica a partir da administração oral de um preparado que contenha o mesmo princípio ativo.
3. Bioequivalência: Demonstração de equivalência farmacêutica entre produtos apresentados sob a mesma forma farmacêutica, contendo idêntica composição qualitativa e quantitativa de princípio(s) ativo(s), e que tenham comparável biodisponibilidade (quantidade e velocidade de absorção do fármaco no sistema circulatório), quando estudados sob um mesmo delineamento experimental.
4. Denominação Comum Brasileira (DCB): Denominação do fármaco ou princípio farmacologicamente ativo aprovada pelo órgão federal responsável pela vigilância sanitária.
5. Denominação Comum Internacional (DCI): Denominação do fármaco ou princípio farmacologicamente ativo recomendada pela Organização Mundial da Saúde.
6. Fármacos de Liberação Imediata: Efeito do composto químico liberado após ingestão.
7. Fármacos de Liberação Modificada: Incorporação de novos efeitos ao fármaco por uso de tecnologias farmacêuticas.
8. Food and Drug Administration (FDA): Agência de proteção do consumidor do governo dos Estados Unidos que monitora dispositivos médicos, alimentos e drogas.
9. Forma Farmacêutica: Estado final de apresentação que os princípios ativos farmacêuticos possuem após uma ou mais operações farmacêuticas executadas com a adição de excipientes

apropriados ou sem a adição de excipientes, a fim de facilitar a sua utilização e obter o efeito terapêutico desejado, com características apropriadas a uma determinada via de administração.

10. Formulação: Relação quantitativa dos farmoquímicos que compõe um medicamento.

11. Matriz Biológica: Material individualizado de origem biológica que pode ser amostrado e processado de maneira reprodutiva.

12. Medicamento de referência: Produto inovador registrado no órgão federal responsável pela vigilância sanitária e comercializado no País, cuja eficácia, segurança e qualidade foram comprovadas cientificamente junto a ANVISA por ocasião do registro.

13. Parâmetros Farmacocinéticos: Variáveis respostas dos experimentos farmacêuticos.

14. REBLAS: Denominação comum dos laboratórios habilitados que integram a Rede Brasileira de Laboratórios Analíticos em Saúde.

15. Testes Analíticos: Testes de qualidade preconizados pela Farmacopéia Brasileira. Na ausência da monografia na Farmacopéia Brasileira, serão aceitos outros compêndios oficiais reconhecidos pelo Ministério da Saúde. Na ausência de monografia farmacopêica, serão aceitos testes analíticos acompanhados de sua validação segundo GUIA DE VALIDAÇÃO DE TESTES ANALÍTICOS.